



Defective Sensor Identification for WSNs involving Generic Local Outlier Detection Tests

Wenjie Li, Francesca Bassi, Davide Dardari, Michel Kieffer, Gianni Pasolini

► To cite this version:

Wenjie Li, Francesca Bassi, Davide Dardari, Michel Kieffer, Gianni Pasolini. Defective Sensor Identification for WSNs involving Generic Local Outlier Detection Tests. IEEE Transactions on Signal and Information Processing over Networks, 2016, 2 (1), pp.29-48. 10.1109/TSIPN.2016.2516821 . hal-01260533

HAL Id: hal-01260533

<https://hal.science/hal-01260533>

Submitted on 26 Jan 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Defective Sensor Identification for WSNs involving Generic Local Outlier Detection Tests

Wenjie Li¹, Francesca Bassi^{1,2}, Davide Dardari³, Michel Kieffer^{1,4}, and Gianni Pasolini³

Abstract

The behavior of a wireless sensor network dedicated to distributed estimation tasks may be significantly altered by the presence of nodes whose sensors are defective and produce erroneous measurements. This paper proposes and analyzes the performance of two distributed algorithms to help each node in determining whether it is equipped with a defective sensor. A node first collects data from its neighborhood, processes them to decide, using some generic local outlier detection test, whether these data contain outliers and broadcasts the result. Then, it determines the status of its own sensor using its result and those received from neighboring nodes. A single-decision and an iterative algorithm for defective sensor detection are proposed. Bounds on the performance of the single-decision algorithm are derived. A theoretical analysis of the probability of error and of the equilibrium of the iterative algorithm is provided for a wide class of local outlier detection tests. The trade-off between false alarm probability and detection probability is characterized theoretically and by simulation. MAC-layer issues, as well as the effect of packet losses are accounted for.

I. INTRODUCTION

The behavior of a wireless sensor network (WSN) [2] dedicated to distributed estimation tasks may be significantly altered by the presence of nodes whose sensors produce *outliers*, *i.e.* abnormal measurements which cannot be justified by the mere effects of sensing noise [3]. It is thus very important to detect such nodes efficiently (with low communication and computation costs), ideally in a distributed way. Many efforts have been considered in this direction in the last years see [4], [5], [6] and references therein. As far as the distributed fault detection (DFD) is concerned, in particular, many different techniques have been proposed in the literature (see the survey paper [5] for a detailed classification). Some techniques rely only on data coming from the direct neighborhood of each node [7], others also use data coming from farther nodes [8], [9]. Some techniques are model-based and rely on a parametric model of the system observed by the sensors to identify outliers [10], [11], whereas model-free techniques mainly exploit statistical properties of the measurements, such as spatial and temporal correlation. For example,

This work has been partly supported by the NEWCOM# NoE. Part of the results in this paper has been presented at IEEE International Conference on Communications (ICC), London, UK, June 9-12, 2015 [1].

¹W. Li, F. Bassi, and M. Kieffer are with Laboratoire des Signaux et Systèmes (L2S, UMR CNRS 8506) CNRS-CentraleSupélec-Université Paris-Sud, 3, rue Joliot Curie 91192 Gif-sur-Yvette, France `firstname.name@lss.supelec.fr`

²F. Bassi is also with ESME-Sudria, 94200 Ivry-sur-Seine, France

³D. Dardari and G. Pasolini are with CNIT, DEI, University of Bologna, Italy. `firstname.name@unibo.it`

⁴M. Kieffer is also with LTCI Telecom ParisTech, 75013 Paris, France and Institut Universitaire de France, 75005 Paris, France

the algorithm in [7] allows each node to estimate its own status by comparing its local measurements with those at neighboring nodes. In a first phase only a tendency status (*good*, *faulty*, *likely good*, or *likely faulty*) is determined. In the second phase the tendency status of neighbors are collected and associated to their measurements to obtain a more reliable assessment. In [12] the local test is based on the comparison between the local measurement and the median of the measurements of the neighbors. In [13], a modified three-sigma edit test is proposed to identify nodes producing measurements with very high variance. Iterative algorithms are proposed in [14], [15], where the weighted-median and the weighted average criterion are considered, respectively. In both cases the local test weights the measurements of the neighbors by the confidence level obtained from the previous detection rounds, under the assumption of permanent node failure. In [16], the algorithm uses time redundancy to tolerate transient faults in sensing and in communication. The adaptive algorithm proposed in [17] adjusts the decision threshold at each round to improve its detection accuracy. In [9], a generic outlier identification function is assumed to be available at each node. The results provided by this function at each node as well as measurements are exchanged to allow the whole network to identify all outliers in a given dataset. In [8], a distributed Bayesian approach is proposed to detect the outliers in a large set of data collected by a WSN.

In general the availability of more data at a given node facilitates the detection and identification of outliers, at the price of higher communication costs and detection delays. An interesting approach to reduce the number of tests is group testing (GT), which is a statistical method allowing to identify a small number of outliers within a large set of data by performing only a limited number of elementary tests, referred to as local outlier detection tests (LODTs), on data subsets [18]. The elementary test is only able to determine whether outliers are present in the considered set. A distributed GT-based algorithm extending the results of centralized GT [19] is proposed in [20], under the assumption of vanishing ratio of defective sensors. The solution provided in [20] suffers however from a significant drawback in the communication overhead necessary for node indexing and for the dissemination of the test results from the cluster heads to the rest of the WSN.

Despite the significant number of papers on this topic, there are still open issues to be addressed, as pointed out in [5], [6]. It is well known, in this regard, that WSNs are usually constrained in terms of memory, computational power and energy sources. Most of the traditional fault detection schemes, conceived for wired networks, have paid little attention to these issues. They are usually memory and energy eager, and are often high demanding in terms of communication effort and computational complexity. It is certainly true, on the one hand, that the exchange of messages and their processing are the only means of fault diagnosis; on the other hand, however, outlier detection schemes targeted to WSNs must be lightweight and impose a limited communication cost.

Moreover, the literature is mainly focused on new or improved LODTs, which are validated considering specific case studies through simulation, thus without determining *a priori* the relationship between the parameters of the algorithm (*e.g.*, detection threshold) and the performance. At the authors' best knowledge, no attempts are present to widen the perspective, generalizing the DFD methodology and providing equilibrium or stability analysis of iterative DFD algorithms. The latter are very important issues as the knowledge of the conditions under which the algorithm reaches the equilibrium and is stable, prevents the system to fail in an unpredictable way.

In order to address these issues, in this paper we propose and analyze two low-delay, low-complexity DFD algorithms inspired by the GT technique: a single-decision one and an iterative one, the second being better suited to WSNs with a large proportion of defective sensors. In both cases we decouple the LODT, that provides only an indication (true, false) of the existence of outliers in the neighborhood, from the final decision on the node status (good, faulty). The proposed algorithms consist, in fact, of two stages, namely *local outlier detection* using a LODT, and *decision*. In the first stage, each node collects data from its neighborhood, performs the LODT to determine whether outliers are present among the collected data, and broadcasts the result. Then, each node estimates the status of its sensor in the decision stage, on the basis of the outcomes of the LODTs performed in the neighborhood.

The two-stage nature of the algorithm makes the decision phase agnostic for what the LODT is regarded. This is particularly relevant because in this paper we provide the quite mild assumptions the LODT should satisfy in order to analytically prove the existence of an equilibrium condition. Moreover, the same analytical framework allows to predict the performance of the algorithm and tune the value of its parameters. This aspect is worth special attention, as most of the DFD algorithms proposed by the literature rely on some decision thresholds, whose values must be properly defined according to the given specification (in terms of *detection probability* and *false alarm probability*, for instance). As pointed out in [4], [5], this is still an open issue and one of the shortcomings of many outlier detection strategies, whose thresholds are chosen according to heuristic guidelines or via simulation in some specific scenarios.

Some preliminary results of our work have been presented in [1]. In this extended version, the DFD algorithms have been modified to cope with channel access issues and transmission impairments resulting in packet losses. Theoretical results in [1] involve some conjectures, whereas here sufficient conditions to be satisfied by the LODT have been established to ensure the existence of an equilibrium of the DFD algorithms. These results are general, since no specific form of the LODT need to be considered. We show that the LODT is only required to determine whether outliers are present in a set of data, without necessarily being able to identify which data are erroneous, thus widening the range of applicable LODTs to those requiring a few measurements from neighboring nodes. Such LODTs are easily accessible, *e.g.*, in the context of bounded-error estimation [21], [22] and may provide good results even with as few as two or three measurements available to perform the test. This aspect is crucial in the WSNs context because it allows to relieve the network traffic and prolong its lifetime. Let us stress, in this regard, that most of traditional outlier detection tests (based on the median [14] or the mean [15] of measurements, for instance) are not very efficient when only few measurements are available. Furthermore, the low complexity of such LODTs is well-suited to the stringent resource (memory, computational power, etc.) constraints that usually affect sensor nodes.

Finally, the trade-off between false alarm probability and detection probability is characterized theoretically and by simulation for some outlier models in realistic channel conditions.

Summarizing, the main novelties of this paper are:

- Proposed a low-complexity DFD algorithm based on very generic LODTs that have the potential to work with very few measurements;

- Found the mild conditions to be satisfied by the LODT and by the density of nodes in the network to ensure that the iterative DFD algorithm leads to an equilibrium;
- Characterized analytically the performance in realistic channel conditions, related to packet losses and to channel access issues at the MAC layer, the latter being seldom considered in other papers.

The rest of the paper is organized as follows. Section II presents the system model and basic assumptions. Section III describes the two variants of the proposed DFD algorithm. Section IV develops the theoretical analysis of the proposed algorithms. Section V presents some outlier models and a LODT. Section VI provides some numerical results and Section VII concludes this paper.

II. NOTATIONS AND SYSTEM MODEL

A. Network model

Consider an infinite plane where nodes, **equipped with one sensor each**, are uniformly and independently deployed, with spatial density ρ , according to a 2D **homogeneous** Poisson point process (PPP) [23]. Let \mathcal{S} denote the set of nodes. Each sensor observes some physical phenomenon and produces measurements perturbed by random noise. **The noise samples at different nodes, or corrupting different measurements at the same node, are assumed statistically independent.** A sensor is defective if it produces outliers. Several examples of sensing noise and outlier models are introduced in Section V-A. The sets of nodes equipped with defective and non-defective (good) sensors are denoted by \mathcal{D} and \mathcal{G} respectively, with $\mathcal{D} \cup \mathcal{G} = \mathcal{S}$. The spatial densities of nodes with defective and good sensors are ρ_d and ρ_g , respectively, with $\rho = \rho_d + \rho_g$. Let θ_i denote the status (defective or **good**) of an arbitrary sensor node $i \in \mathcal{S}$. One has $\theta_i = 1$ if $i \in \mathcal{D}$ and $\theta_i = 0$ if $i \in \mathcal{G}$. In this work, it is assumed that θ_i remains constant in the time interval during which the DFD is performed.

Assume that any pair of nodes (i, j) can communicate only if $r_{i,j} \leq R_0$, where $r_{i,j}$ is their physical distance and R_0 is the communication range, dependent on the transmission power, transmitter and receiver characteristics, and propagation scenario. Define $\mathcal{N}_i = \{j \in \mathcal{S} \mid 0 < r_{i,j} \leq R_0\}$ as the set of the neighbors of Node i . The number of nodes in \mathcal{N}_i follows a Poisson distribution

$$\mathbb{P}\{|\mathcal{N}_i| = n\} = \frac{\bar{\mu}^n}{n!} \exp(-\bar{\mu}), \quad (1)$$

with $\bar{\mu} = \rho\pi R_0^2$, see [24]. The probability of having n_1 nodes with good sensors and $n_2 = n - n_1$ nodes with defective sensors in \mathcal{N}_i is

$$\begin{aligned} \mathbb{P}\{|\mathcal{N}_i \cap \mathcal{G}| = n_1, |\mathcal{N}_i \cap \mathcal{D}| = n_2\} &= \frac{\bar{\mu}_g^{n_1} \bar{\mu}_d^{n_2}}{n_1! n_2!} \exp(-(\bar{\mu}_g + \bar{\mu}_d)), \\ &= P(\bar{\mu}_g, \bar{\mu}_d, n_1, n_2), \end{aligned} \quad (2)$$

where $\bar{\mu}_g = \rho_g \pi R_0^2$ and $\bar{\mu}_d = \rho_d \pi R_0^2$.

When nodes try to broadcast messages to their neighbors, we assume that some collision avoidance mechanism is put at work to limit packet losses. As a consequence, in a given finite time interval, all nodes are not necessarily able to access the channel to broadcast their message. Moreover, a message broadcast by some node is not necessarily received by all its neighbors due to transmission errors, residual packet collisions, shadowing, etc. Packet integrity

is assumed to be determined with some CRC or checksum. Corrupted packets are considered as lost. To model all these phenomena, we assume that in the communication interval Δt , Node i only receives the messages transmitted by a subset $\mathcal{N}'_i \subseteq \mathcal{N}_i$ of its neighbors. The probability of having n nodes in \mathcal{N}'_i still follows a Poisson distribution

$$\mathbb{P}\{|\mathcal{N}'_i| = n\} = \frac{\mu^n}{n!} \exp(-\mu), \quad (3)$$

where $\mu = \rho\pi R_0^2(1 - \varepsilon(\Delta t))$ is the average number of nodes from which Node i receives a message during the transmission interval, see [24]. We assume that $\varepsilon(\Delta t)$ is a decreasing function of Δt accounting for all the above mentioned channel access impairments. This is reasonable, because as Δt increases more nodes have the opportunity to access the channel.

B. Local outlier detection test

Consider a generic set of nodes $\mathcal{A} \subseteq \mathcal{S}$ and define

$$\varphi(\mathcal{A}) = \begin{cases} 1, & \text{if } \mathcal{A} \cap \mathcal{D} \neq \emptyset, \\ 0, & \text{otherwise,} \end{cases} \quad (4)$$

as the indicator function for the presence of nodes with defective sensors in \mathcal{A} . At a given time instant, let the random variable $M_i \in \mathbb{R}^\eta$ denote the η -dimensional data provided by the sensor of Node i , $i \in \mathcal{A}$. M_i may be a scalar or vector measurement, may contain a measurement and the value of some regressor in the case of system models linear in their parameters, or may contain measurements and experimental conditions, for general nonlinear system models.

The data provided by all nodes in \mathcal{A} are gathered in $\mathbf{M}_{\mathcal{A}} = [M_i]_{i \in \mathcal{A}} \in \mathbb{R}^{\eta|\mathcal{A}|}$. For a given realization $\mathbf{m}_{\mathcal{A}} \in \mathbb{R}^{\eta|\mathcal{A}|}$ of $\mathbf{M}_{\mathcal{A}}$, denote $T(\mathbf{m}_{\mathcal{A}})$ the outcome of some LODT

$$T(\mathbf{m}_{\mathcal{A}}) = \begin{cases} 0, & \text{if no outlier is detected from } \mathbf{m}_{\mathcal{A}}, \\ 1, & \text{otherwise.} \end{cases} \quad (5)$$

Even if $\varphi(\mathcal{A}) = 1$, i.e., at least one node with a defective sensor belongs to \mathcal{A} , the noise characteristics of the defective sensors may produce realizations $\mathbf{m}_{\mathcal{A}}$ that do not allow for the detection of the outliers. Thus, one introduces the probability of detection of the LODT

$$q_D(\mathbf{M}_{\mathcal{A}}) = \mathbb{P}\{T(\mathbf{M}_{\mathcal{A}}) = 1 | \varphi(\mathcal{A}) = 1\}. \quad (6)$$

Similarly, one considers the probability of false alarm of the LODT

$$q_{FA}(\mathbf{M}_{\mathcal{A}}) = \mathbb{P}\{T(\mathbf{M}_{\mathcal{A}}) = 1 | \varphi(\mathcal{A}) = 0\}. \quad (7)$$

In this work we consider only LODTs satisfying the following properties.

1. Let $n_g = |\mathcal{A} \cap \mathcal{G}|$ and $n_d = |\mathcal{A} \cap \mathcal{D}|$, then

$$q_{FA}(\mathbf{M}_{\mathcal{A}}) = q_{FA}(n_g), \quad q_D(\mathbf{M}_{\mathcal{A}}) = q_D(n_g, n_d). \quad (8)$$

2. Let k be an arbitrary node in \mathcal{A} , then

$$\mathbb{P}\{T(\mathbf{M}_{\mathcal{A}\setminus\{k\}}) = 1 \mid T(\mathbf{M}_{\mathcal{A}}) = 0, \varphi(\mathcal{A}) = 0\} = 0. \quad (9)$$

$$\mathbb{P}\{T(\mathbf{M}_{\mathcal{A}\setminus\{k\}}) = 1 \mid T(\mathbf{M}_{\mathcal{A}}) = 0, \varphi(\mathcal{A}) = 1\} = 0. \quad (10)$$

In Property 1, we consider only the class of LODTs $T(\mathbf{M}_{\mathcal{A}})$ processing the random data vector $\mathbf{M}_{\mathcal{A}}$ in a way that disregards the knowledge of the identity of the node producing the data. For this reason, (6) and (7) depend only on the number of sensors in each status belonging to \mathcal{A} . In practice, the LODT takes a realization $\mathbf{m}_{\mathcal{A}}$ as input and does not need to know n_g and n_d . The notations in (8) indicate the dependence of the probabilities q_{FA} and q_D in n_g and n_d that will be exploited in the analysis of the proposed DFD algorithms involving the LODT outcomes.

Property 2 implies that if no outlier is detected testing the whole vector $\mathbf{m}_{\mathcal{A}}$, then no outlier will be detected testing any sub-vector of $\mathbf{m}_{\mathcal{A}}$. LODTs satisfying Properties 1 and 2 are characterized by probabilities of false alarm and of detection depending on n_g and n_d as described by the following lemma.

Lemma 1. *The following inequalities hold for any LODT $T(\mathbf{M}_{\mathcal{A}})$ satisfying Properties 1 and 2:*

$$q_{\text{FA}}(n_g) \leq q_{\text{FA}}(n_g + 1), \quad (11)$$

$$q_D(n_g, n_d) \leq q_D(n_g + 1, n_d), \quad (12)$$

$$q_D(n_g, n_d) \leq q_D(n_g, n_d + 1). \quad (13)$$

The proof of Lemma 1 is given in Appendix A. Lemma 1 implies that a LODT satisfying Properties 1 and 2 is more likely to detect an outlier when the number of data involved in the test increases. Examples of LODTs satisfying Properties 1 and 2 are provided in Section V-B.

Properties 1 and 2 correspond to mild sufficient conditions a LODT has to satisfy to ensure the existence of an equilibrium of Algorithm 2 introduced in Section III-B. LODTs, which do not satisfy these properties, may also be considered, but the behavior of Algorithm 2 cannot be analyzed theoretically in such cases.

III. DFD ALGORITHM

This section proposes two DFD algorithms (a single-decision variant and an iterative variant). In the single-decision variant, several measurement rounds and LODTs are performed by each node. At the end of the last round, a *single* decision concerning the status of its sensor is taken locally by each node. In the iterative variant, a decision is taken at each round and may be updated at the next one. As will be seen in Section VI, the most appropriate variant depends on the ratio $\bar{\mu}_d/\bar{\mu}$ of defective sensors in the network. In what follows, we assume that the network topology does not change within a round of both algorithms.

A. Single-decision DFD algorithm

The proposed single-decision DFD algorithm is described in Algorithm 1. The local outlier detection stage consists of two successive phases (Phase I and Phase II), alternating during L rounds. Then, the decision stage consists of a final decision phase, whose outcome is an estimate $\hat{\theta}_i$ of the status θ_i of the sensor of Node i .

In the ℓ -th round, during Phase I, Node i collects fresh data transmitted by its neighborhood \mathcal{N}_i . A subset containing a proportion ζ of randomly chosen data among the received ones is then used with Node i fresh privacy data to perform the LODT described in Section II-B. Define $\mathcal{U}_i = \mathcal{N}_i \cup \{i\}$ and let $\mathcal{N}_i^{(\ell,1)} \subset \mathcal{N}_i$ be the subset of nodes from which Node i received data (excluding itself) during Phase I of Round ℓ . Consider

$$\mathcal{V}_i^{(\ell,1)} = S_\zeta \left(\mathcal{N}_i^{(\ell,1)} \right) \cup \{i\} \quad (14)$$

the set of node indices whose data are involved in the LODT performed by Node i at Round ℓ . The function S_ζ selects a random subset of nodes in $\mathcal{N}_i^{(\ell,1)}$ such that

$$\left| S_\zeta \left(\mathcal{N}_i^{(\ell,1)} \right) \right| = \left\lceil \zeta \left| \mathcal{N}_i^{(\ell,1)} \right| \right\rceil, \quad (15)$$

where $\left| \mathcal{N}_i^{(\ell,1)} \right|$ is the cardinal number of $\mathcal{N}_i^{(\ell,1)}$ and $\lceil \cdot \rceil$ denotes upwards rounding. The role of ζ is illustrated in Example 3 and further discussed in Section IV-C. The LODT outcome is denoted $Y_i^{(\ell)} = T(\mathbf{M}_{\mathcal{V}_i^{(\ell,1)}})$, which, for a given realization $\mathbf{M}_{\mathcal{V}_i^{(\ell,1)}} = \mathbf{m}_{\mathcal{V}_i^{(\ell,1)}}$, provides $y_i^{(\ell)} = T(\mathbf{m}_{\mathcal{V}_i^{(\ell,1)}})$. Then, during Phase II, Node i tries to broadcast $(y_i^{(\ell)}, \mathcal{V}_i^{(\ell,1)})$ to indicate the other nodes which data were involved in the LODT it just performed. At the end of Phase II, Node i has received pairs $(y_j^{(\ell)}, \mathcal{V}_j^{(\ell,1)})$ from a subset of nodes $\mathcal{V}_i^{(\ell,II)} \subseteq \mathcal{U}_i$. Denote

$$\mathcal{B}_i^{(\ell)} = \left\{ j \in \mathcal{V}_i^{(\ell,II)} \text{ such that } i \in \mathcal{V}_j^{(\ell,1)} \right\}. \quad (16)$$

Node i then adds to a first counter z_i all $y_j^{(\ell)}$ s such that $j \in \mathcal{B}_i^{(\ell)}$. Thus only LODT outcomes involving the data produced by Node i are added to z_i . Node i also accumulates in a second counter n_i the cardinal number $\left| \mathcal{B}_i^{(\ell)} \right|$ of $\mathcal{B}_i^{(\ell)}$. This counter represents the number of LODT outcomes received by Node i in which its data were involved (including its own LODT outcomes). At the end of the L rounds the value z_i/n_i is the statistics on which the decision is taken in the final decision phase.

The decision (19) can result in both false alarm, with probability P_{FA} , and non-detection, with probability $P_{\text{ND}} = 1 - P_{\text{D}}$, where P_{D} is the detection probability. The value of ζ and of the threshold γ affect the trade-off between P_{D} and P_{FA} , and have to be adjusted to meet the targeted performance (see Appendix B for more details). For a fixed value of ζ , when γ is close to one, P_{D} may be low. On the other hand, P_{FA} may be high for values of γ close to zero. Increasing L provides a better averaging effect in (19), which reduces the variance of $\hat{\theta}_i$. Nevertheless, L cannot be taken too large, to preserve the hypothesis that the status θ_i does not vary during the whole DFD procedure.

As will be seen in Section VI, the single-decision DFD algorithm performs well when $\bar{\mu}_{\text{d}} \ll \bar{\mu}_{\text{g}}$. However, P_{FA} rapidly increases with $\bar{\mu}_{\text{d}}/\bar{\mu}$. Consider the case where $\bar{\mu}_{\text{d}}$ is large enough so that every non-defective sensor node has at least one defective sensor node in its neighborhood. When $\zeta = 1$, all measurements from neighbors are involved in the LODT of Node i . As a consequence, for all sensors $i \in \mathcal{S}$, the LODT outcome will be $y_i^{(\ell)} = 1$ with a high probability. Non-defective sensors are then frequently diagnosed as defective. Reducing ζ increases the chance for Node i to get an outlier-free subset of data on which the LODT can be performed and thus reduces P_{FA} . The idea of selecting only a subset of data to perform the LODT is reminiscent to GT [19].

Now, when $\bar{\mu}$ is small, *i.e.*, the degree of connectivity of the network is low, it may be useful to collect measurements from the h -hop neighborhood of each node and to broadcast test outcomes to the same h -hop

Algorithm 1 Single-decision DFD

1) Initialize $\hat{\theta}_i^{(0)} = 0$, $z_i = 0$, and $n_i = 0$ for all $i \in \mathcal{S}$.

2) For each round $1 \leq \ell \leq L$:

- Phase I, **lasting** $\triangle t_I$: Node i tries to broadcast a packet containing its local data $m_i^{(\ell)}$, receives the data produced by the nodes in \mathcal{N}_i , **randomly selects a subset** $S_\zeta \left(\mathcal{N}_i^{(\ell,1)} \right)$ **of received data** and performs the test (5) with outcome

$$y_i^{(\ell)} = T \left(\mathbf{m}_{\mathcal{V}_i^{(\ell,1)}} \right); \quad (17)$$

- Phase II, **lasting** $\triangle t_{II}$: each node broadcasts $\left(y_i^{(\ell)}, \mathcal{V}_i^{(\ell,1)} \right)$ generated in Phase I and updates z_i **and** n_i as follows

$$\begin{cases} z_i = z_i + \sum_{j \in \mathcal{B}_i^{(\ell)}} y_j^{(\ell)}, \\ n_i = n_i + \left| \mathcal{B}_i^{(\ell)} \right|. \end{cases} \quad (18)$$

3) After L rounds:

- Decision phase: each node i estimates the status θ_i of its sensor

$$\hat{\theta}_i = \begin{cases} 1 \text{ (defective)} & \text{if } z_i/n_i \geq \gamma, \\ 0 \text{ (non defective)} & \text{otherwise,} \end{cases} \quad (19)$$

where γ is some threshold, $0 < \gamma \leq 1$.

neighborhood. Appendix C describes the updates required in Algorithm 1 to handle DFD with multi-hop data collection and LODT result dissemination.

Examples 2 and 3 illustrate the behavior of the single decision DFD algorithm in various situations.

Example 2. Figure 1 depicts a **first example WSN**, **where no channel access issues nor collisions are considered**, thus $\varepsilon = 0$. Let k be the only node equipped with a defective sensor. Assume that both q_{FA} defined in (7) and q_{ND} defined in (6) are negligible: all the sensors in \mathcal{U}_k successfully detect outliers. Moreover consider $L = 1$ (unique round) and take $\gamma = 0.7$ and $\zeta = 1$. At the end of Phase II, $\mathcal{B}_i^{(1)} = \mathcal{U}_i$, $i = 1, \dots, k$. **Moreover**, $z_k^{(1)} = |\mathcal{U}_k| = 7$, $z_1^{(1)} = |\mathcal{U}_k \cap \mathcal{U}_1| = 4$, $z_2^{(1)} = |\mathcal{U}_k \cap \mathcal{U}_2| = 1$, and $z_3^{(1)} = |\mathcal{U}_k \cap \mathcal{U}_3| = 0$. Since $z_k^{(1)} / |\mathcal{U}_k| = 1 > \gamma$, $z_1^{(1)} / |\mathcal{U}_1| \approx 0.67 < \gamma$, $z_2^{(1)} / |\mathcal{U}_2| = 0.25 < \gamma$, and $z_3^{(1)} / |\mathcal{U}_3| = 0 < \gamma$, only Node k determines its sensor as defective, according to (19), while Nodes 1, 2, and 3 diagnose their **own** sensor as non-defective. **All decisions are thus correct.**

Example 3. Figure 2 represents a second example of a WSN with a minimum degree of connectivity. Node 2 is equipped with a defective sensor. With the same assumptions as in Example 2, considering $L = 1$, $\varepsilon = 0$, $\gamma = 0.7$, and $\zeta = 1$, Node 1 and 2 will determine themselves as defective. Allowing two-hop data collection and the LODT outcome dissemination worsens the situation since all nodes will have the outlier produced by Node 2 in their collected data. As a consequence, all LODTs will detect the presence of an outlier, and all nodes will determine

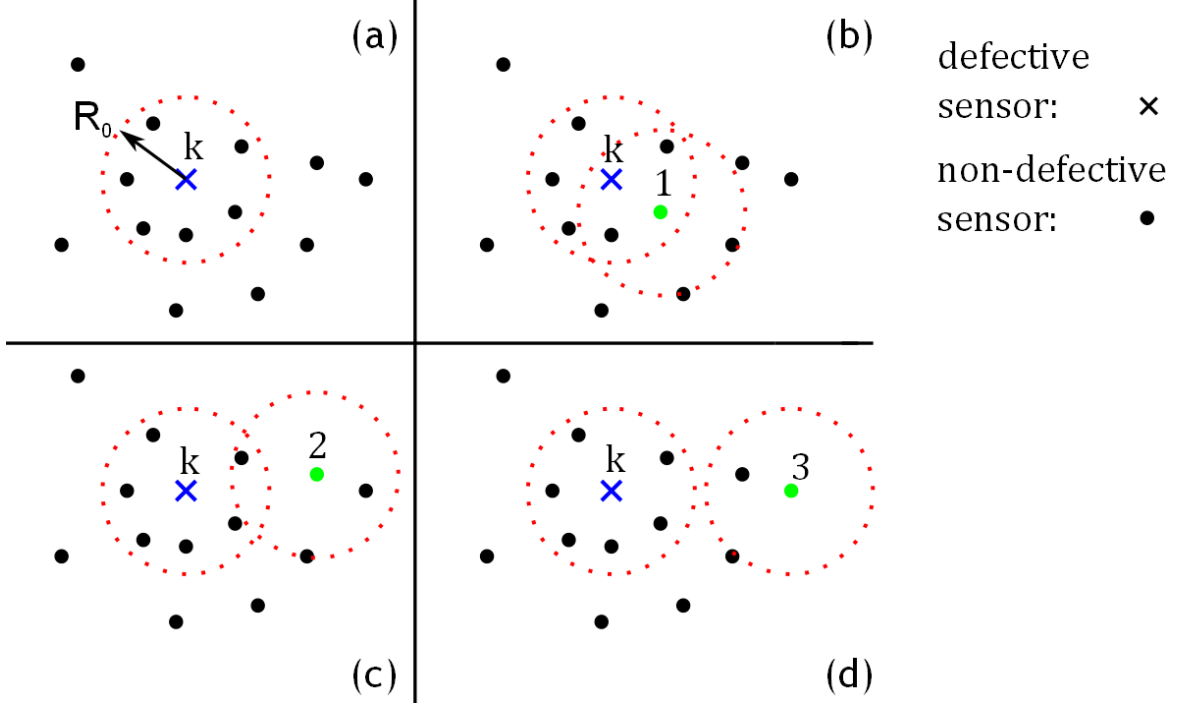


Fig. 1. Example where k is the only node with a defective sensor, (a) shows that $z_k^{(1)} = |\mathcal{U}_k| = 7$. In (b), Node 1 belongs to \mathcal{U}_k and $z_1^{(1)} = 4$. In (c), Node 2 is not in \mathcal{U}_k but $r_{2,k} \leq 2R_0$, which results in $z_2^{(1)} = 1$. In (d), Node 3 has a distance $r_{3,k} > 2R_0$, so $z_3^{(1)} = 0$.



Fig. 2. Example where Node 2 is the only defective sensor.

themselves as defective. This is why it is necessary to perform the LODT only on a subset of data. Consider again a two-hop data collection with $\zeta = 0.5$. In Phase I, Nodes 1 and 4 receive data from Nodes 2 and 3. Their LODTs will provide $y_1^{(1)} = 0$ with a probability 0.5 and $y_4^{(1)} = 0$ with a probability 0.5. Node 2 always uses its own data and produces $y_2^{(1)} = 1$. Node 3 receives data from Nodes 1, 2, and 4. It randomly selects $\lceil 0.5 \times 3 \rceil = 2$ data out of these 3 data and provides thus $y_3^{(1)} = 0$ with a probability $2/3$. In Phase II, Node 1 will receive LODT outcomes from Nodes 2 and 3. Only outcomes such that $y_i^{(1)} = 1$ and $\{1\} \subset \mathcal{V}_i^{(1,1)}$ contribute to z_i . Moreover, all outcomes with $\{1\} \subset \mathcal{V}_i^{(1,1)}$ contribute to n_1 . The third line of Table I provides $\Pr(\{j\} \subset \mathcal{V}_i^{(1,1)})$, $i, j = 1, \dots, 4$. For example, $\Pr(\{1\} \subset \mathcal{V}_1^{(1,1)}) = 1$, $\Pr(\{1\} \subset \mathcal{V}_2^{(1,1)}) = 2/3$. One deduces for example

$$E(N_1) = \sum_{i=1}^4 \Pr(\{1\} \subset \mathcal{V}_i^{(1,1)}) = 7/3.$$

The fourth line of Table I provides $\Pr(\{j, 2\} \subset \mathcal{V}_i^{(1,1)})$, $i, j = 1, \dots, 4$. This helps to evaluate $E(Z_i)$. For example,

$$E(Z_1) = \sum_{i=1}^4 \Pr(\{1, 2\} \subset \mathcal{V}_i^{(1,1)}) = 11/6.$$

Node i	1	2	3	4
$\Pr(Y_i^{(1)} = 1)$	0.5	1	0.66	0.5
$\Pr(\{j\} \subset \mathcal{V}_i^{(1,I)})$	(1, 1/2, 1/2, 0)	(2/3, 1, 2/3, 2/3)	(2/3, 2/3, 1, 2/3)	(0, 1/2, 1/2, 1)
$\Pr(\{j, 2\} \subset \mathcal{V}_i^{(1,I)})$	(1/2, 1/2, 0, 0)	(2/3, 1, 2/3, 2/3)	(1/3, 2/3, 2/3, 1/3)	(0, 1/2, 0, 1/2)
$E(Z_i)$	3/2	8/3	4/3	3/2
$E(N_i)$	7/3	8/3	8/3	7/3
$E(Z_i)/E(N_i)$	9/14	1	1/2	9/14

TABLE I
BEHAVIOR OF THE SINGLE DECISION DFD FOR EXAMPLE 3

The last line of Table I provides $E(Z_i)/E(N_i)$. One observes that a proper selection of $\gamma > 9/14$ allows in average to detect the node producing an outlier, while avoiding false alarms. One also observes that $E(Z_i)/E(N_i)$ is larger for Nodes 1 and 4 which are connected (at one or two hops) to fewer good nodes than Node 3. This illustrates the fact that the level of protection against a false alarm is related to the number of connections to good nodes.

Example 3 illustrates the importance of the choices of the values of γ and ζ . Some insights are provided for their tuning in Section IV-C and Appendix B.

B. Iterative DFD algorithm

To improve the single-decision DFD algorithm for increasing $\bar{\mu}_d/\bar{\mu}$, an iterative variant of Algorithm 1 is described in Algorithm 2. The iterative algorithm is composed of two parts, of duration L_1 and L_2 rounds, respectively. In each round of the first part (Lines 1 to 4 of Algorithm 2) **nodes collect data from their neighborhood, perform a LODT, and** are allowed to take temporary decisions $\hat{\theta}_i^{(\ell)}$ about the status of their sensor (*i.e.*, **a single round consists of both the local outlier detection and the decision stages**). The temporary decision at Round $\ell - 1$ affects the set of sensor data tested by Node i at Round ℓ , *i.e.*, **the nodes with $\hat{\theta}_i^{(\ell-1)} = 1$ remain silent during Round ℓ** . **The set of active neighbors of Node i is then $\hat{\mathcal{U}}_i^{(\ell)} = \{j \in \mathcal{U}_i \text{ s.t. } \hat{\theta}_j^{(\ell-1)} = 0\}$ and one has $\mathcal{V}_i^{(\ell,I)} \subseteq \hat{\mathcal{U}}_i^{(\ell)}$ and $\mathcal{V}_i^{(\ell,II)} \subseteq \hat{\mathcal{U}}_i^{(\ell)}$** . Any temporary decision can be updated during the next rounds of the first part. **This is why z_i and n_i are reinitialized at the beginning of Phase I in each round of this first part**. Finally, Algorithm 1 is performed during the second part (Lines 6 and 7 of Algorithm 2) involving only nodes with sensors deemed as non-defective.

The core procedure of the first part is described in Lines 2 and 3. One has $\hat{\theta}_i^{(\ell)} = 1$ if Node i and all neighbors **which have used the data of Node i** detect an outlier. Note that (20) is similar to (19) with $L = 1$ and $\gamma = 1$. This choice has been considered in the first part of Algorithm 2 to minimize the number of nodes with sensors wrongly deemed as defective, while having a reasonable probability of detection P_D . This allows then to work in the second part of Algorithm 2 with a smaller value of $\bar{\mu}_d/\bar{\mu}$. This is especially important when $\bar{\mu}_d \geq 1$: a large amount of non-defective sensors may be diagnosed as defective even with $\gamma = 1$. In Line 2, if $\hat{\theta}_i^{(\ell-1)} = 1$, Node i stops broadcasting its data in Round ℓ . Nevertheless, Node i still performs the LODT: this gives it a chance to become active again at Round $\ell + 1$, in case it erroneously considered its sensor as defective at Round ℓ .

Algorithm 2 Iterative DFD

- 1) Set $\ell = 1$; $\widehat{\theta}_i^{(0)} = 0$ for all $i \in \mathcal{S}$.
 - 2) **Phase I**, lasting Δt_I : Set $z_i = n_i = 0$.
 - a) If $\widehat{\theta}_i^{(\ell-1)} = 0$, Node i tries to broadcast a packet containing its local data $m_i^{(\ell)}$, receives the data produced by the nodes in $\mathcal{N}_i^{(\ell,I)} \subset \mathcal{N}_i$, randomly selects a subset $S_\zeta(\mathcal{N}_i^{(\ell,I)})$ of received data and performs the test (5) with outcome $y_i^{(\ell)} = T(\mathbf{m}_{\mathcal{V}_i^{(\ell,I)}})$;
 - b) Else Node i is silent, *i.e.*, it does not broadcast its data but receives data from its neighbors.
 - 3) **Phase II**, lasting Δt_{II} :
 - a) If $\widehat{\theta}_i^{(\ell-1)} = 0$, Node i tries broadcast $(y_i^{(\ell)}, \mathcal{V}_i^{(\ell,I)})$, evaluates z_i and n_i , and performs the decision

$$\widehat{\theta}_i^{(\ell)} = \begin{cases} 1, & \text{if } z_i/n_i = 1, \\ 0, & \text{otherwise.} \end{cases} \quad (20)$$
 - b) Else, Node i performs the decision $\widehat{\theta}_i^{(\ell)} = y_i^{(\ell)}$.
 - 4) $\ell = \ell + 1$.
 - 5) If $\ell \leq L_1$, go to 2.
 - 6) After round L_1 :
 - each node i such that $\widehat{\theta}_i^{(L_1)} = 0$ sets $z_i = n_i = 0$.
 - each node i such that $\widehat{\theta}_i^{(L_1)} = 1$ is determined as defective. It stops broadcasting its data. It does not participate in the single-decision DFD during the following rounds.
 - 7) Single-decision DFD is performed during L_2 rounds with threshold γ .
-

IV. ANALYSIS OF THE PROPOSED DFD ALGORITHMS

In Section IV-A we characterize analytically the probabilities of detection P_D and of false alarm P_{FA} of the proposed DFD algorithms, when the decision is taken after collecting a single data by each node in the neighborhood. These represent P_D and P_{FA} of the single-decision DFD algorithm for $L = 1$, or the probability of detection and false alarm on the temporary decision of the first stage of the iterative algorithm. Conditions for the existence of an equilibrium of the iterative algorithm are also evaluated in Section IV-C. An analysis of the traffic generated by both algorithms is then performed in Sections IV-B and IV-E.

A. P_D and P_{FA} for a single round of the single-decision DFD algorithm

To lighten the notations, the round index is omitted in this section, *e.g.*, $Y_i^{(\ell)}$, $\mathcal{V}_i^{(\ell,I)}$ and $\mathcal{V}_i^{(\ell,II)}$ are replaced by Y_i , $\mathcal{V}_i^{(I)}$ and $\mathcal{V}_i^{(II)}$ respectively. Moreover, $\varepsilon(\Delta t_I)$ and $\varepsilon(\Delta t_{II})$ are simplified to ε_I and ε_{II} respectively. Assume that Node i successfully accesses the channel during Phase I with probability α_1 . One has $\alpha_1 \geq 1 - \varepsilon_I$ since ε_I also accounts for channel impairments, such as packet collisions.

In order to perform the analysis of P_D and P_{FA} for a decision involving a single round of data collection from the neighborhood of Node i , one has to characterize the probability $\mathbb{P}\{Y_i = 1\}$ of the LODT (5). This probability

depends on the numbers $N_g^{(I)} = |\mathcal{V}_i^{(I)} \cap \mathcal{G}|$ and $N_d^{(I)} = |\mathcal{V}_i^{(I)} \cap \mathcal{D}|$ of nodes equipped with good and defective sensors in $\mathcal{V}_i^{(I)}$. Define the function $h(n_g, n_d)$ as the probability a LODT yields $Y_i = 1$, conditioned on $N_g^{(I)} = n_g$ and $N_d^{(I)} = n_d$, i.e.,

$$\begin{aligned} h(n_g, n_d) &= \mathbb{P}\{Y_i = 1 \mid N_g^{(I)} = n_g, N_d^{(I)} = n_d\} \\ &= \begin{cases} q_{\text{FA}}(n_g), & \text{if } n_d = 0, \\ q_{\text{D}}(n_g, n_d), & \text{if } n_d \neq 0. \end{cases} \end{aligned} \quad (21)$$

Now, $\mathbb{P}\{Y_i = 1 \mid \theta_i = 0\}$ can be expressed as a function of $\mu_g^{(I)} = \alpha_2 \zeta \bar{\mu}_g$ and $\mu_d^{(I)} = \alpha_2 \zeta \bar{\mu}_d$, respectively, where $\alpha_2 = (1 - \varepsilon_1)$ indicates the probability Node i successfully received a packet from one of its neighbors and ζ accounts for the probability a node has to select in its LODT the data produced by a neighboring node. The rounding effects have been neglected. One has thus

$$\begin{aligned} f_1(\mu_g^{(I)}, \mu_d^{(I)}) &= \mathbb{P}\{Y_i = 1 \mid \theta_i = 0\} \\ &\stackrel{(a)}{=} \sum_{n_g=1}^{\infty} \sum_{n_d=0}^{\infty} \mathbb{P}\{Y_i = 1 \mid N_g^{(I)} = n_g, N_d^{(I)} = n_d, \theta_i = 0\} \\ &\quad \cdot \mathbb{P}\{N_g^{(I)} = n_g, N_d^{(I)} = n_d \mid \theta_i = 0\} \\ &\stackrel{(b)}{=} \sum_{n_g=1}^{\infty} \sum_{n_d=0}^{\infty} h(n_g, n_d) \cdot P(\mu_g^{(I)}, \mu_d^{(I)}, n_g - 1, n_d), \end{aligned} \quad (22)$$

where P is the function defined in (2). Moreover, (a) is due to the fact that if $i \in \mathcal{V}_i^{(I)}$ and $\theta_i = 0$, then $N_g^{(I)} \geq 1$, and (b) comes from Property 1, which states that q_{FA} and q_{D} depend only on the number of good and defective sensors. Similarly, one may introduce

$$\begin{aligned} f_2(\mu_g^{(I)}, \mu_d^{(I)}) &= \mathbb{P}\{Y_i = 1 \mid \theta_i = 1\} \\ &= \sum_{n_g=0}^{\infty} \sum_{n_d=1}^{\infty} h(n_g, n_d) \cdot P(\mu_g^{(I)}, \mu_d^{(I)}, n_g, n_d - 1). \end{aligned} \quad (23)$$

In order to characterize P_{FA} , introduce $N_g^{(\text{II})} = |\mathcal{B}_i \cap \mathcal{G}|$, $N_d^{(\text{II})} = |\mathcal{B}_i \cap \mathcal{D}|$ and the conditional false alarm event

$$\mathcal{E}_{i, \text{FA}}^{n_g, n_d} = \left\{ \frac{\sum_{j \in \mathcal{B}_i} Y_j}{n_g + n_d} \geq \gamma \mid N_g^{(\text{II})} = n_g, N_d^{(\text{II})} = n_d, \theta_i = 0 \right\} \quad (24)$$

representing, according to (20), the situation where Node i with non-defective sensor diagnoses it as defective, knowing that $N_g^{(\text{II})} = n_g \geq 1$ and $N_d^{(\text{II})} = n_d \geq 0$. Introducing

$$\tau_{\text{FA}}(n_g, n_d) = \mathbb{P}\left\{ \mathcal{E}_{i, \text{FA}}^{n_g, n_d} \right\}, \quad (25)$$

then P_{FA} can be expressed as

$$P_{\text{FA}} = \sum_{n_g=1}^{\infty} \sum_{n_d=0}^{\infty} \tau_{\text{FA}}(n_g, n_d) \mathbb{P}\left\{ N_g^{(\text{II})} = n_g, N_d^{(\text{II})} = n_d \mid \theta_i = 0 \right\}. \quad (26)$$

Two situations need to be considered. If Node i fails to access the channel, then $\mathcal{B}_i = \{i\}$, which means that $N_g^{(\text{II})} = 1$ and $N_d^{(\text{II})} = 0$. Otherwise $N_g^{(\text{II})} - 1$ and $N_d^{(\text{II})}$ follow Poisson distributions. To evaluate their averages $\mu_g^{(\text{II})}$

and $\mu_d^{(\text{II})}$, one has to consider the probability $\alpha_3 = (1 - \varepsilon_I)(1 - \varepsilon_{\text{II}})/\alpha_1$ that Node i received a packet in Phase II from one of its neighbors, assuming that Node i actually accessed the channel in Phase I. Moreover, knowing that such packet comes from a neighbor of Node i , it contains a decision involving the data of Node i with a probability equal to ζ . Consequently, $\mu_g^{(\text{II})} = \alpha_3 \zeta \bar{\mu}_g$ and $\mu_d^{(\text{II})} = \alpha_3 \zeta \bar{\mu}_d$. Then

$$P_{\text{FA}} = \alpha_1 \tilde{P}_{\text{FA}}(\mu_g^{(\text{II})}, \mu_d^{(\text{II})}) + (1 - \alpha_1) f_1(\mu_g^{(\text{I})}, \mu_d^{(\text{I})}) \quad (27)$$

where

$$\begin{aligned} \tilde{P}_{\text{FA}}(\mu_g^{(\text{II})}, \mu_d^{(\text{II})}) &= \sum_{n_g=1}^{\infty} \sum_{n_d=0}^{\infty} \tau_{\text{FA}}(n_g, n_d) \\ &\quad \cdot P(\mu_g^{(\text{II})}, \mu_d^{(\text{II})}, n_g - 1, n_d). \end{aligned} \quad (28)$$

Similarly, consider Node i with defective sensor. The *conditional defective sensor detection event* is

$$\mathcal{E}_{i,\text{D}}^{n_g, n_d} = \left\{ \frac{\sum_{j \in \mathcal{B}_i} Y_j}{n_g + n_d} \geq \gamma \mid N_g^{(\text{II})} = n_g, N_d^{(\text{II})} = n_d, \theta_i = 1 \right\}, \quad (29)$$

where $n_g \geq 0$ and $n_d \geq 1$. Introducing

$$\tau_{\text{D}}(n_g, n_d) = \mathbb{P}\{\mathcal{E}_{i,\text{D}}^{n_g, n_d}\}, \quad (30)$$

then P_{D} can be expressed as

$$P_{\text{D}} = \alpha_1 \tilde{P}_{\text{D}}(\mu_g^{(\text{II})}, \mu_d^{(\text{II})}) + (1 - \alpha_1) f_2(\mu_g^{(\text{I})}, \mu_d^{(\text{I})}), \quad (31)$$

where

$$\begin{aligned} \tilde{P}_{\text{D}}(\mu_g^{(\text{II})}, \mu_d^{(\text{II})}) &= \sum_{n_g=0}^{\infty} \sum_{n_d=1}^{\infty} \tau_{\text{D}}(n_g, n_d) \\ &\quad \cdot P(\mu_g^{(\text{II})}, \mu_d^{(\text{II})}, n_g, n_d - 1). \end{aligned} \quad (32)$$

Lemma 4. P_{FA} and P_{D} are decreasing functions of γ .

Proof: The monotonicity of P_{FA} and P_{D} with respect to γ comes from the fact that if $\gamma_1 > \gamma_2$, $\sum_{j \in \mathcal{B}_i} Y_j \geq \gamma_1 (n_g + n_d)$ implies $\sum_{j \in \mathcal{B}_i} Y_j \geq \gamma_2 (n_g + n_d)$. ■

In what follows, we assume that γ is fixed, and express P_{FA} and P_{D} as functions of $\mu_g^{(\text{II})}$ and $\mu_d^{(\text{II})}$. To find a closed-form expression for P_{FA} and P_{D} , one has to evaluate the pmf of $(\sum_{j \in \mathcal{B}_i} Y_j \mid N_g^{(\text{II})}, N_d^{(\text{II})}, \theta_i)$. Consider now the event

$$\mathcal{Y}_{j,\text{D}}^{n_g, n_d} = \left\{ Y_j = 1 \mid j \in \mathcal{B}_i, N_g^{(\text{II})} = n_g, N_d^{(\text{II})} = n_d, \theta_i = 1 \right\},$$

with $j \in \mathcal{B}_i$. For any $k \in \mathcal{B}_i$ with $j \neq k$, $\mathcal{Y}_{j,\text{D}}^{n_g, n_d}$ and $\mathcal{Y}_{k,\text{D}}^{n_g, n_d}$ are dependent. Their dependency comes from the fact that in general $\mathcal{N}_j \cap \mathcal{N}_k \neq \emptyset$. The pmf of $\sum_{j \in \mathcal{B}_i} Y_j$ is thus quite difficult to evaluate, since the dependency between the Y_j s is not explicit. For this reason, only upper bounds of P_{FA} and P_{D} are derived in what follows.

Lemma 5. Consider an arbitrary γ with $0 < \gamma \leq 1$, then for all $\mu_g > 0$ and $\mu_d \geq 0$, one has

$$P_{\text{FA}} \left(\mu_g^{(\text{II})}, \mu_d^{(\text{II})} \right) \leq f_1 (4\bar{\mu}_g, 4\bar{\mu}_d), \quad (33)$$

$$P_{\text{D}} \left(\mu_g^{(\text{II})}, \mu_d^{(\text{II})} \right) \leq f_2 (4\bar{\mu}_g, 4\bar{\mu}_d). \quad (34)$$

The proof of Lemma 5 is given in Appendix E. The upper bounds in Lemma 5 are tight only if $\gamma \rightarrow 0$ but loose if $\gamma = 1$. Since $\gamma = 1$ in the first stage of the iterative algorithm, tighter upper bounds for $P_{\text{FA}} \left(\mu_g^{(\text{II})}, \mu_d^{(\text{II})} \right)$ and $P_{\text{D}} \left(\mu_g^{(\text{II})}, \mu_d^{(\text{II})} \right)$ are needed in this situation.

Lemma 6. If $\gamma = 1$, one has

$$P_{\text{FA}} \left(\mu_g^{(\text{II})}, \mu_d^{(\text{II})} \right) \leq f_1 \left(\mu_g^{(\text{II})}, \mu_d^{(\text{II})} \right), \quad (35)$$

$$P_{\text{D}} \left(\mu_g^{(\text{II})}, \mu_d^{(\text{II})} \right) \leq f_2 \left(\mu_g^{(\text{II})}, \mu_d^{(\text{II})} \right). \quad (36)$$

Proof: By definition of $\mathcal{E}_{i,\text{FA}}^{n_g, n_d}$ in (24), if $\gamma = 1$, one has

$$\begin{aligned} & \tau_{\text{FA}} (n_g, n_d) \\ &= \mathbb{P} \left\{ \bigcap_{j \in \mathcal{B}_i} \{Y_j = 1\} \mid \theta_i = 0, N_g = n_g, N_d = n_d \right\} \\ &\leq \mathbb{P} \left\{ Y_i = 1 \mid \theta_i = 0, N_g = n_g, N_d = n_d \right\} \\ &= h(n_g, n_d). \end{aligned} \quad (37)$$

Then (35) can be obtained from (22), (28), and (37). Equation (36) can be obtained similarly. \blacksquare

B. Traffic generated by the single-decision DFD algorithm

This analysis assumes that a single-hop data collection and LODT result dissemination is employed.

During Phase I of each round ℓ , the data $m_i^{(\ell)}$ collected by each node has to be transmitted, with its identifier, usually already present in the packet header. This type of information typically fits a single IEEE 802.15.4 payload [2].

During Phase II, $(y_i^{(\ell)}, \mathcal{V}_i^{(\ell,1)})$ has to be broadcast. One single bit is needed for $y_i^{(\ell)}$. The transmission of $\mathcal{V}_i^{(\ell,1)}$ is necessary if the network topology is dynamic or in presence of packet losses. This information requires $(\bar{\mu} + 1) \lceil \log_2 |\mathcal{S}| \rceil$ bits in average, assuming that each node can be identified with an index of $\lceil \log_2 |\mathcal{S}| \rceil$ bits only. In a static network, nodes usually know the index of their neighbors. Thus, only the indexes of neighbors which do not participate to the LODT of Node i need to be transmitted. These nodes belong to $\mathcal{U}_i \setminus \mathcal{V}_i^{(\ell,1)}$. In this case, the required average number of bits boils down to $(1 - \alpha_2 \zeta) \bar{\mu} \lceil \log_2 |\mathcal{S}| \rceil$.

Without the knowledge of $\mathcal{V}_i^{(\ell,1)}$, the algorithm can still be performed with somewhat degraded performance, as will be shown in Section VI-C.

$$\begin{cases} \bar{\mu}_{00}^{(\ell+1)} = F_1 \left(\bar{\mu}_{00}^{(\ell)}, \bar{\mu}_{10}^{(\ell)} \right) = \alpha_1 \bar{\mu}_{00}^{(\ell)} \left(f_1 \left(\alpha_2 \zeta \bar{\mu}_{00}^{(\ell)}, \alpha_2 \zeta \bar{\mu}_{10}^{(\ell)} \right) - \tilde{P}_{\text{FA}} \left(\alpha_3 \zeta \bar{\mu}_{00}^{(\ell)}, \alpha_3 \zeta \bar{\mu}_{10}^{(\ell)} \right) \right) + \bar{\mu}_{\text{g}} \left(1 - f_1 \left(\alpha_2 \zeta \bar{\mu}_{00}^{(\ell)}, \alpha_2 \zeta \bar{\mu}_{10}^{(\ell)} \right) \right), \\ \bar{\mu}_{10}^{(\ell+1)} = F_2 \left(\bar{\mu}_{00}^{(\ell)}, \bar{\mu}_{10}^{(\ell)} \right) = \alpha_1 \bar{\mu}_{10}^{(\ell)} \left(f_2 \left(\alpha_2 \zeta \bar{\mu}_{00}^{(\ell)}, \alpha_2 \zeta \bar{\mu}_{10}^{(\ell)} \right) - \tilde{P}_{\text{D}} \left(\alpha_3 \zeta \bar{\mu}_{00}^{(\ell)}, \alpha_3 \zeta \bar{\mu}_{10}^{(\ell)} \right) \right) + \bar{\mu}_{\text{d}} \left(1 - f_2 \left(\alpha_2 \zeta \bar{\mu}_{00}^{(\ell)}, \alpha_2 \zeta \bar{\mu}_{10}^{(\ell)} \right) \right). \end{cases} \quad (41)$$

C. Equilibrium of the iterative algorithm

Let the pair $(\theta_i, \hat{\theta}_i)$ denote the *state* of the sensor of Node i , where θ_i is its actual status and $\hat{\theta}_i$ is its estimated status. Among the four possible states, (0, 0) and (1, 1) are states resulting from a correct decision, (0, 1) corresponds to a false alarm, and (1, 0) corresponds to a non-detection. Let $\bar{\mu}_{\theta_i, \hat{\theta}_i}^{(\ell)}$ denote the density of sensors in the state $(\theta_i^{(\ell)}, \hat{\theta}_i^{(\ell)})$ at the beginning of round ℓ . The aim of this section is to characterize the evolution of $\bar{\mu}_{\theta_i, \hat{\theta}_i}^{(\ell)}$ to determine whether the iterative algorithm converges to a steady state.

Before the first round of the iterative algorithm, one has $\bar{\mu}_{00}^{(0)} = \bar{\mu}_{\text{g}}$, $\bar{\mu}_{10}^{(0)} = \bar{\mu}_{\text{d}}$, and $\bar{\mu}_{01}^{(0)} = \bar{\mu}_{11}^{(0)} = 0$. Note that at any round ℓ , $\bar{\mu}_{00}^{(\ell)} + \bar{\mu}_{01}^{(\ell)} = \bar{\mu}_{\text{g}}$ and $\bar{\mu}_{10}^{(\ell)} + \bar{\mu}_{11}^{(\ell)} = \bar{\mu}_{\text{d}}$. Since the actual status of a sensor is assumed constant during the rounds of the DFD algorithm, the only possible transitions are between states (0, 0) and (0, 1) and between states (1, 0) and (1, 1). The evolution of the densities during round ℓ is given by

$$\begin{pmatrix} \bar{\mu}_{00}^{(\ell+1)} \\ \bar{\mu}_{01}^{(\ell+1)} \\ \bar{\mu}_{10}^{(\ell+1)} \\ \bar{\mu}_{11}^{(\ell+1)} \end{pmatrix} = \begin{pmatrix} P_{0,00}^{(\ell)} & P_{0,10}^{(\ell)} & 0 & 0 \\ P_{0,01}^{(\ell)} & P_{0,11}^{(\ell)} & 0 & 0 \\ 0 & 0 & P_{1,00}^{(\ell)} & P_{1,10}^{(\ell)} \\ 0 & 0 & P_{1,01}^{(\ell)} & P_{1,11}^{(\ell)} \end{pmatrix} \begin{pmatrix} \bar{\mu}_{00}^{(\ell)} \\ \bar{\mu}_{01}^{(\ell)} \\ \bar{\mu}_{10}^{(\ell)} \\ \bar{\mu}_{11}^{(\ell)} \end{pmatrix}, \quad (38)$$

where $P_{a,bc}^{(\ell)} = \mathbb{P} \left\{ \hat{\theta}_i^{(\ell)} = c \mid \hat{\theta}_i^{(\ell-1)} = b, \theta_i = a \right\}$ is the transition probability from (a, b) to (a, c) , for any $a, b, c \in \{0, 1\}$. The sensors considered as defective in the previous iteration are silent, so all the transition probabilities are functions of $\bar{\mu}_{00}^{(\ell)}$ and $\bar{\mu}_{10}^{(\ell)}$. For different values of $\hat{\theta}_i^{(\ell-1)}$, $\hat{\theta}_i^{(\ell)}$ is obtained according to Line 2 of Algorithm 2. More precisely, if $\hat{\theta}_i^{(\ell-1)} = 0$, one applies (20). According to (27) and (31), one has

$$\begin{cases} P_{0,01}^{(\ell)} = \alpha_1 \tilde{P}_{\text{FA}} \left(\mu_{00}^{(\ell, \text{I})}, \mu_{10}^{(\ell, \text{I})} \right) + (1 - \alpha_1) f_1 \left(\mu_{00}^{(\ell, \text{I})}, \mu_{10}^{(\ell, \text{I})} \right), \\ P_{0,00}^{(\ell)} = 1 - P_{0,01}^{(\ell)}, \\ P_{1,01}^{(\ell)} = \alpha_1 \tilde{P}_{\text{D}} \left(\mu_{00}^{(\ell, \text{II})}, \mu_{10}^{(\ell, \text{II})} \right) + (1 - \alpha_1) f_2 \left(\mu_{00}^{(\ell, \text{I})}, \mu_{10}^{(\ell, \text{I})} \right), \\ P_{1,00}^{(\ell)} = 1 - P_{1,01}^{(\ell)}, \end{cases} \quad (39)$$

where $\mu_{\theta_i, \hat{\theta}_i}^{(\ell, \text{I})} = \alpha_2 \zeta \bar{\mu}_{\theta_i, \hat{\theta}_i}^{(\ell)}$ and $\mu_{\theta_i, \hat{\theta}_i}^{(\ell, \text{II})} = \alpha_3 \zeta \bar{\mu}_{\theta_i, \hat{\theta}_i}^{(\ell)}$. If $\hat{\theta}_i^{(\ell-1)} = 1$, one has $\hat{\theta}_i^{(\ell)} = y_i^{(\ell)}$ and

$$\begin{cases} P_{0,11}^{(\ell)} = f_1 \left(\mu_{00}^{(\ell, \text{I})}, \mu_{10}^{(\ell, \text{I})} \right), & P_{0,10}^{(\ell)} = 1 - P_{0,11}^{(\ell)}, \\ P_{1,11}^{(\ell)} = f_2 \left(\mu_{00}^{(\ell, \text{I})}, \mu_{10}^{(\ell, \text{I})} \right), & P_{1,10}^{(\ell)} = 1 - P_{1,11}^{(\ell)}. \end{cases} \quad (40)$$

From (38-40), one thus obtains the reduced non-linear state equation (41) (at the top of the next page), describing the evolution of $\bar{\mu}_{00}^{(\ell)}$ and $\bar{\mu}_{10}^{(\ell)}$ in the iterative DFD algorithm.

Let μ_{00}^* and μ_{10}^* be values at equilibrium of $\bar{\mu}_{00}^{(\ell)}$ and $\bar{\mu}_{10}^{(\ell)}$, respectively. From (41), one deduces that μ_{00}^* and μ_{10}^* should satisfy

$$\frac{\alpha_1 \mu_{00}^* \tilde{P}_{\text{FA}}(\alpha_3 \zeta \mu_{00}^*, \alpha_3 \zeta \mu_{10}^*)}{\bar{\mu}_g - \mu_{00}^*} = 1 - \frac{\bar{\mu}_g - \alpha_1 \mu_{00}^*}{\bar{\mu}_g - \mu_{00}^*} f_1(\alpha_2 \zeta \mu_{00}^*, \alpha_2 \zeta \mu_{10}^*), \quad (42)$$

and

$$\frac{\alpha_1 \mu_{10}^* \tilde{P}_{\text{D}}(\alpha_3 \zeta \mu_{00}^*, \alpha_3 \zeta \mu_{10}^*)}{\bar{\mu}_d - \mu_{10}^*} = 1 - \frac{\bar{\mu}_d - \alpha_1 \mu_{10}^*}{\bar{\mu}_d - \mu_{10}^*} f_2(\alpha_2 \zeta \mu_{00}^*, \alpha_2 \zeta \mu_{10}^*). \quad (43)$$

The existence property of the equilibrium point depend on the monotonicity of f_1 , f_2 , \tilde{P}_{FA} , and \tilde{P}_{D} .

Lemma 7. $f_1(\mu_{00}, \mu_{10})$ and $f_2(\mu_{00}, \mu_{10})$ are non-decreasing functions of μ_{00} and μ_{10} .

Proof: One has

$$\begin{aligned} \frac{\partial f_1}{\partial \mu_{00}} &= \sum_{n_g=1}^{\infty} \sum_{n_d=0}^{\infty} \frac{\mu_{10}^{n_d} \mu_{00}^{n_g-1}}{n_d! (n_g-1)!} \exp(-\mu_{10} - \mu_{00}) \\ &\quad \cdot (h(n_g+1, n_d) - h(n_g, n_d)), \end{aligned} \quad (44)$$

as shown in Appendix D. From Lemma 1, one has $h(n_g, n_d+1) \geq h(n_g, n_d)$ and then $\frac{\partial f_1}{\partial \mu_{00}} \geq 0$. Similarly, one has $\frac{\partial f_1}{\partial \mu_{10}} \geq 0$, $\frac{\partial f_2}{\partial \mu_{00}} \geq 0$, and $\frac{\partial f_2}{\partial \mu_{10}} \geq 0$. ■

The following lemma provides sufficient conditions to have monotone left hand-side expressions in (42) and (43).

Lemma 8. Assume that the LODT (5) satisfies Properties 1 and 2. Assume also that for some reference Node i and some node with random index $K \in \mathcal{N}_i$ with $\theta_i = \theta_K = 0$ and $Y_i = 1$, the LODT is such that

$$\mathbb{P}\{Y_K = 1 \mid Y_i = 1, \theta_i = \theta_K = 0, K \in \mathcal{B}_i\} > 1 - \frac{4}{\alpha_3 \zeta \bar{\mu}_g}, \quad (45)$$

$$q_{\text{D}}(0, 2) > 1 - \frac{4}{\alpha_3 \zeta \bar{\mu}_d}, \quad (46)$$

then, for any μ_{10} ,

$$g_{\text{FA}}(\mu_{00}, \mu_{10}) = \frac{\mu_{00} \tilde{P}_{\text{FA}}(\mu_{00}, \mu_{10})}{\alpha_3 \zeta \bar{\mu}_g - \mu_{00}} \quad (47)$$

is an increasing function of μ_{00} over $[0, \bar{\mu}_g]$ and for any μ_{00} ,

$$g_{\text{D}}(\mu_{00}, \mu_{10}) = \frac{\mu_{10} \tilde{P}_{\text{D}}(\mu_{00}, \mu_{10})}{\alpha_3 \zeta \bar{\mu}_d - \mu_{10}} \quad (48)$$

is an increasing function of μ_{10} over $[0, \bar{\mu}_d]$.

The proof of Lemma 8 is in Appendix F.

Lemma 8 allows us to propose a sufficient condition on the LODT, so that existence and uniqueness of the equilibrium point can be guaranteed.

Proposition 9. Assume that the considered LODT satisfies Properties 1 and 2, as well as (45) and (46), then (42-43) admits a solution, and an equilibrium of (41) exists.

Proof: We study first the behavior of the right-hand side of (42) for a fixed value of μ_{10} . As shown in Section IV-A, both $\frac{\bar{\mu}_g - \alpha_1 \mu_{00}}{\bar{\mu}_g - \mu_{00}}$ and f_1 are positive and non-decreasing function of μ_{00} , then

$$s_1(\mu_{00}, \mu_{10}) = 1 - \frac{\bar{\mu}_g - \alpha_1 \mu_{00}}{\bar{\mu}_g - \mu_{00}} f_1(\alpha_2 \zeta \mu_{00}, \alpha_2 \zeta \mu_{10}) \quad (49)$$

is a non-increasing function of μ_{00} . If (45) is satisfied, the left-hand side of (42), $\alpha_1 g_{\text{FA}}(\alpha_3 \zeta \mu_{00}, \alpha_3 \zeta \mu_{10})$, is a strictly increasing function of μ_{00} over $[0, \bar{\mu}_g]$. Moreover

$$\alpha_1 g_{\text{FA}}(0, \alpha_3 \zeta \mu_{10}) = 0 < 1 - f_1(0, \mu_{10}),$$

and

$$\lim_{\mu_{00} \rightarrow \bar{\mu}_g} \alpha_1 g_{\text{FA}}(\alpha_3 \zeta \mu_{00}, \alpha_3 \zeta \mu_{10}) \gg \lim_{\mu_{00} \rightarrow \bar{\mu}_g} s_1(\mu_{00}, \mu_{10})$$

since $\lim_{\mu_{00} \rightarrow \bar{\mu}_g} g_{\text{FA}}(\mu_{00}, \mu_{10}) \rightarrow \infty$ and $s_1(\mu_{00}, \mu_{10}) \leq 1$. Therefore, there exists a unique intersection of the graphs of functions $\alpha_1 g_{\text{FA}}(\alpha_3 \zeta \mu_{00}, \alpha_3 \zeta \mu_{10})$ and $s_1(\mu_{00}, \mu_{10})$ as μ_{00} goes from 0 to $\bar{\mu}_g$, for any fixed $\mu_{10} \in [0, \bar{\mu}_d]$. The value μ_{00}^+ of μ_{00} at which they intersect is thus a function of μ_{10} . One may write

$$\mu_{00}^+ = \beta_1(\mu_{10}),$$

where $\beta_1 : [0, \bar{\mu}_d] \rightarrow [0, \bar{\mu}_g]$. The continuity of β_1 is deduced from that of s_1 and g_{FA} .

Similarly, one may show that if (46) is satisfied, for any fixed μ_{00} , there exists a unique intersection of the graphs of $\alpha_1 g_{\text{D}}(\alpha_3 \zeta \mu_{00}, \alpha_3 \zeta \mu_{10})$ and

$$s_2(\mu_{00}, \mu_{10}) = 1 - \frac{\bar{\mu}_d - \alpha_1 \mu_{10}}{\bar{\mu}_d - \mu_{10}} f_2(\alpha_2 \zeta \mu_{00}, \alpha_2 \zeta \mu_{10}) \quad (50)$$

as μ_{10} varies from 0 to $\bar{\mu}_d$. The value μ_{10}^+ of μ_{10} at which both graphs intersect is thus a function of μ_{00} . One may write

$$\mu_{10}^+ = \beta_2(\mu_{00}),$$

where $\beta_2 : [0, \bar{\mu}_g] \rightarrow [0, \bar{\mu}_d]$. The continuity of β_2 is deduced from that of f_2 and g_{D} . Then, the equation

$$\mu_{00}^* = \beta_1(\beta_2(\mu_{00}^*))$$

admits at least a solution from Brouwer's fixed-point theorem, since $\beta_1 \circ \beta_2 : [0, \bar{\mu}_g] \rightarrow [0, \bar{\mu}_g]$ is continuous. In this way, Proposition 9 is proved. \blacksquare

The unicity of the equilibrium requires additional monotonicity conditions on α and β , which are not easy to establish for a generic LODT.

If $\alpha_3 \zeta \bar{\mu}_g < 4$ and $\alpha_3 \zeta \bar{\mu}_d < 4$, one has $1 - 4/(\alpha_3 \zeta \bar{\mu}_g) < 0$ and $1 - 4/(\alpha_3 \zeta \bar{\mu}_d) < 0$. Then an equilibrium exists for any type of LODT. In practice, $\bar{\mu}_d$ is usually small and (46) is easily satisfied. When $\alpha_3 \zeta \bar{\mu}_g \geq 4$, (45) imposes some constraint on the LODT: considering two non-defective nodes i and k , the conditional probability $\mathbb{P}\{Y_k = 1 \mid Y_j = 1\}$ should be larger than $1 - 4/(\alpha_3 \zeta \bar{\mu}_g)$. The fact that Y_i and Y_k are dependent when $k \in \mathcal{B}_i$ helps to have $\mathbb{P}\{Y_k = 1 \mid Y_j = 1\}$ relatively large. The expression of a lower bound for $\mathbb{P}\{Y_k = 1 \mid Y_j = 1\}$ depends on the type of LODT (Section V-B). In all cases, for large values of $\bar{\mu}_g$, one has to choose small ζ to ensure equilibrium.

The fact that $\bar{\mu}_g$ should not be too large may appear counter-intuitive. To explain this, assume that R_0 is large enough to allow all nodes to communicate directly with all other nodes and that $\mathcal{B}_i = \mathcal{U}_i$. At each node, in presence of a node with defective sensor, all LODTs are likely to yield $Y_i = 1$ for all nodes. All nodes will then stop communicating their data, perform a LODT at the next round with their own data only, be unable to detect an outlier, and finally turn on again. As a consequence, most of the nodes will turn on and off alternately, preventing the DFD algorithm to reach an equilibrium.

A simple way to address this issue is by reducing $\bar{\mu}_g$ and $\bar{\mu}_d$. This may be done with a smaller R_0 , obtained by lowering the power with which nodes are transmitting data. Alternatively, one may use a small ζ , reducing the number of data involved in each LODT.

D. Local asymptotic stability of the iterative algorithm

Consider a small perturbation $\bar{\mu}_{00}^{(\ell)} = \mu_{00}^* + \delta_0^{(\ell)}$ and $\bar{\mu}_{10}^{(\ell)} = \mu_{10}^* + \delta_1^{(\ell)}$ around equilibrium (μ_{00}^*, μ_{10}^*) . One may linearize (41) around equilibrium to get

$$\begin{pmatrix} \delta_0^{(\ell+1)} \\ \delta_1^{(\ell+1)} \end{pmatrix} = \mathbf{A} \begin{pmatrix} \delta_0^{(\ell)} \\ \delta_1^{(\ell)} \end{pmatrix}, \quad (51)$$

where

$$\mathbf{A} = \begin{bmatrix} \frac{\partial F_1}{\partial \bar{\mu}_{00}}(\mu_{00}^*, \mu_{10}^*) & \frac{\partial F_1}{\partial \bar{\mu}_{10}}(\mu_{00}^*, \mu_{10}^*) \\ \frac{\partial F_2}{\partial \bar{\mu}_{00}}(\mu_{00}^*, \mu_{10}^*) & \frac{\partial F_2}{\partial \bar{\mu}_{10}}(\mu_{00}^*, \mu_{10}^*) \end{bmatrix}.$$

The linearized system (51) is asymptotically stable if the eigenvalues of \mathbf{A} are in the unit circle. Providing stability conditions independently of the LODT is difficult. Nevertheless, for a given LODT, the local stability of (41) may be verified numerically. This requires an evaluation of the derivatives of f_1 , f_2 , P_{FA} and P_D with respect to $\bar{\mu}_{00}$ and $\bar{\mu}_{10}$ at equilibrium (Section VI-A).

E. Traffic generated by the iterative DFD algorithm

Compared to the single-decision DFD algorithm, the traffic generated by the iterative DFD algorithm is less heavy. Consider Round ℓ of the iterative algorithm, only the nodes with $\hat{\theta}_i^{(\ell-1)} = 0$ will try to broadcast packets. As discussed in Section IV-C, the proportion of the active nodes is $\kappa^{(\ell)} = \frac{\bar{\mu}_{00}^{(\ell-1)} + \bar{\mu}_{10}^{(\ell-1)}}{\bar{\mu}}$. Therefore, the average number of bits per sensor to perform an L -round iterative algorithm is

- $\sum_{\ell=1}^L \kappa^{(\ell)} (\bar{\mu} + 1) \lceil \log_2 |\mathcal{S}| \rceil$, if the network topology is time-varying.
- $\sum_{\ell=1}^L \kappa^{(\ell)} (1 - \alpha_2 \zeta) \bar{\mu} \lceil \log_2 |\mathcal{S}| \rceil = \sum_{\ell=1}^L (1 - \alpha_2 \zeta) \left(\bar{\mu}_{00}^{(\ell-1)} + \bar{\mu}_{10}^{(\ell-1)} \right) \lceil \log_2 |\mathcal{S}| \rceil$, if the network is static.

V. APPLICATION EXAMPLES

This section presents different outlier models and an example of LODT to be used in the simulations presented in Section VI.

A. Outlier model

Assume that each sensor i gets a noisy observation of the same scalar physical quantity ϕ

$$x_i = \phi + w_i, \quad \forall i \in \mathcal{S}. \quad (52)$$

The components w_i of the measurement noise in (52) are assumed to be realizations of **independent** random variables W_i . Here three different outlier models are considered, with different distributions of the measurement noise. **When successive measurements are taken by the same sensor, we assume that the noise realizations are also independent.**

- *UHV: Uniform distribution with High Variance.* The measurement noise of a good sensor $i \in \mathcal{G}$ is assumed to be uniformly distributed in the interval $[-\Delta, \Delta]$. For a defective sensor $i \in \mathcal{D}$, W_i is also uniformly distributed, but in a larger interval $[-\xi\Delta, \xi\Delta]$ with $\xi > 1$.
- *GHV: Gaussian distribution with High Variance.* The measurement noise follows a zero-mean Gaussian distribution, *i.e.*,

$$W_i \sim \begin{cases} \mathcal{N}(0, \sigma^2), & \text{if } i \in \mathcal{G}, \\ \mathcal{N}(0, (\xi\sigma)^2), & \text{if } i \in \mathcal{D}, \end{cases} \quad (53)$$

where σ is the standard deviation and $\xi > 1$.

- *BGLV: Bias Gaussian distribution with Low Variance.* In this situation, we assume that $W_i \sim \mathcal{N}(E_i, \sigma^2)$ where the bias E_i is also a random variable. If $i \in \mathcal{G}$, then E_i is uniformly distributed in $[-\Delta, \Delta]$. If $i \in \mathcal{D}$, then E_i is uniformly distributed in $[-\xi\Delta, -\Delta] \cup [\Delta, \xi\Delta]$ with $\xi > 1$. Moreover, the value of E_i of each sensor is constant, *i.e.*, E_i does not vary over time.

B. Local outlier detection test

Consider some threshold $\nu \in \mathbb{R}$ and the interval $[m_i] = [m_i - \nu, m_i + \nu]$ of width 2ν centered around each measurement m_i . Consider a set of nodes $\mathcal{A} \subset \mathcal{S}$ and the intersection of all $[m_i]$ s with $i \in \mathcal{A}$

$$[\hat{\phi}(\mathbf{m}_{\mathcal{A}})] = \bigcap_{i \in \mathcal{A}} [m_i]. \quad (54)$$

With and without presence of outliers, one is able to evaluate $\mathbb{P}\left\{\left[\hat{\phi}\right] = \emptyset\right\}$ as a function of σ , ξ , and Δ for the three models introduced in Section V-A. Equation (54) can be used to define a low-complexity LODT

$$T(\mathbf{m}_{\mathcal{A}}) = \begin{cases} 1, & \text{if } [\hat{\phi}(\mathbf{m}_{\mathcal{A}})] = \emptyset, \\ 0, & \text{else.} \end{cases} \quad (55)$$

The following example illustrates the behavior of the LODT in (55) with a small number of measurements.

1. Consider three sensors measuring some constant temperature, *e.g.*, with actual value $t = 20^\circ\text{C}$. Suppose that non-defective sensors have a bounded measurement error, *e.g.*, $\pm 1^\circ\text{C}$. Assume that the local measurement of the first sensor is $t_1 = 19.5^\circ\text{C}$, and that two other sensors provide $t_2 = 20.8^\circ\text{C}$ and $t_3 = 18.2^\circ\text{C}$, respectively. Assuming that there is no defective sensor, and taking into account the bounded measurement noise, one deduces that $t \in \mathbf{t}_1 = [t_1 - 1, t_1 + 1] = [18.5, 20.5]$, $t \in \mathbf{t}_2 = [19.8, 21.8]$, and $t \in \mathbf{t}_3 = [17.2, 19.2]$. **However since**

$\mathbf{t}_1 \cap \mathbf{t}_2 \cap \mathbf{t}_3 = \emptyset$, either the bounds on the measurement noise are too optimistic, or there is at least one outlier. Considering the second hypothesis, it remains difficult to determine which sensor produces an outlier, as $\mathbf{t}_1 \cap \mathbf{t}_2 \neq \emptyset$ and $\mathbf{t}_1 \cap \mathbf{t}_3 \neq \emptyset$.

The test (55) allows to detect the presence of an outlier based on $\mathbf{m}_{\mathcal{A}}$ only, without having to consider the identity of nodes producing each measurement. This test does not give any insight on the nodes with defective sensors. Thus q_{FA} and q_{D} defined in (7) and (6) only depend on n_{g} and n_{d} , where $n_{\text{g}} = |\mathcal{A} \cap \mathcal{G}|$ and $n_{\text{d}} = |\mathcal{A} \cap \mathcal{D}|$. Property 1 is thus satisfied.

$T(\mathbf{m}_{\mathcal{A}}) = 0$ implies $[\hat{\phi}(\mathbf{m}_{\mathcal{A}})] \neq \emptyset$. As consequence, $[\hat{\phi}(\mathbf{m}_{\mathcal{A} \setminus \{j\}})] \neq \emptyset$ for any $j \in \mathcal{A}$, since

$$[\hat{\phi}(\mathbf{m}_{\mathcal{A}})] = \bigcap_{i \in \mathcal{A}} [m_i] \subseteq \bigcap_{i \in \mathcal{A} \setminus \{j\}} [m_i] = [\hat{\phi}(\mathbf{m}_{\mathcal{A} \setminus \{j\}})].$$

Therefore, (10) holds and Property 2 is satisfied as well.

Hereafter we investigate the conditions introduced by Proposition 9.

Lemma 10. *Consider the LODT defined in (55). Consider a reference Node i and a node with random index $K \in \mathcal{N}_i$ such that $\theta_i = \theta_K = 0$, and $Y_i = 1$, then*

$$\mathbb{P}\{Y_K = 1 \mid Y_i = 1, \theta_i = \theta_K = 0, K \in \mathcal{B}_i\} \geq 1 - \frac{\sqrt{3}}{\pi} - \frac{5}{6\pi^2}. \quad (56)$$

See Appendix H for the proof. From (56), one obtains that the condition (45) is $1 - \frac{\sqrt{3}}{\pi} - \frac{5}{6\pi^2} > 1 - \frac{4}{\alpha_3 \zeta \bar{\mu}_{\text{g}}}$ and thus $\alpha_3 \zeta \bar{\mu}_{\text{g}} < \frac{24\pi^2}{5+6\sqrt{3}\pi} \approx 6.3$.

VI. SIMULATION RESULTS

We consider a WSN of 1000 nodes randomly deployed according to a 2D PPP over a square of size 10×10 units, with $\bar{\mu} = 6$. To avoid boundary effects, only nodes in the central square area of size $(10 - 2R_0) \times (10 - 2R_0)$ units are considered in the evaluations of P_{D} and P_{FA} . The parameters of the outlier models and of the LODT are in Table II.

TABLE II
PARAMETERS OF THE OUTLIER MODELS PRESENTED IN SECTION V-A AND OF THE LODT PRESENTED IN SECTION V-B.

	outlier model	outlier detection test
UHV	$\xi = 10$	$\nu = 2\Delta$
GHV	$\xi = 10$	$\nu = 6\sigma$
BGLV	$\xi = 10, \Delta = 10\sigma$	$\nu = 2\Delta$

The values of ν reported in Table II have been chosen in each case via Monte-Carlo simulation of the measurement models for $n_{\text{g}} = 10$ and $n_{\text{d}} = 1$. Since $q_{\text{D}}(n_{\text{g}}, n_{\text{d}})$ and $q_{\text{FA}}(n_{\text{g}})$ are increasing functions of n_{g} , the derived values of q_{D} and q_{FA} are valid upper bounds for all $n_{\text{g}} \in [1, 10]$. Meanwhile, one uses $n_{\text{d}} = 1$ so that the obtained q_{D} is the lower bound for all values of $n_{\text{d}} \geq 1$. Figure 3 represents the evolutions of $q_{\text{D}}(10, 1)$ and $q_{\text{FA}}(10)$ as a function of $\omega = \nu/\Delta$ for the UHV and BGLV outlier models, and $\omega = \nu/(3\sigma)$ for the GHV outlier model. With the UHV

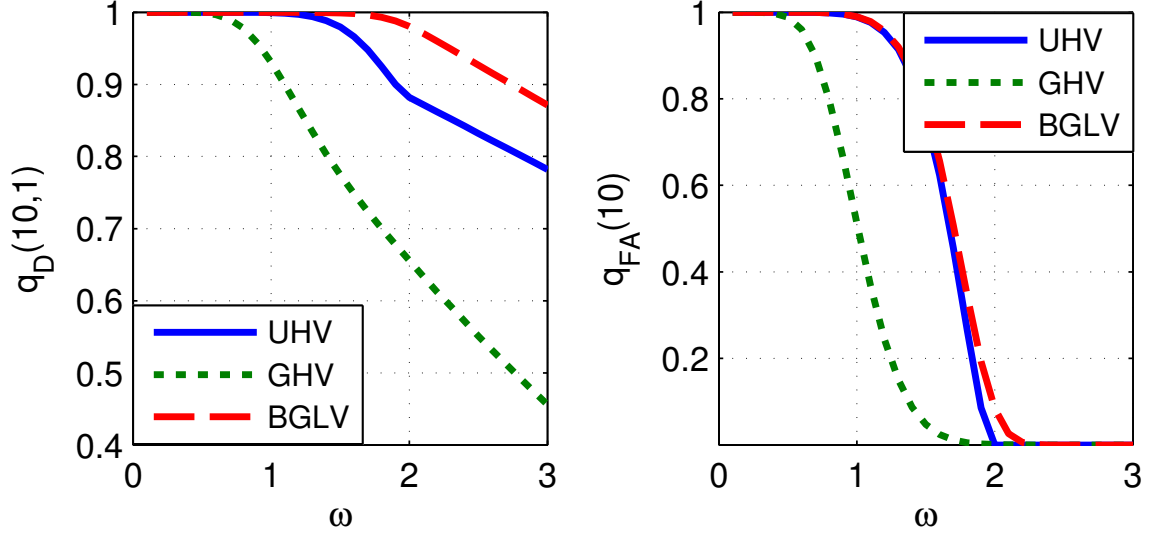


Fig. 3. $q_D(10,1)$ (left) and $q_{FA}(10)$ (right) as functions of $\omega = \nu/\Delta$ for the UHV and BGLV outlier models, and $\omega = \nu/(3\sigma)$ for the GHV outlier model

outlier model, the noise corrupting data produced by a good sensor is bounded in $[-\Delta, \Delta]$, thus $q_{FA} = 0$ if $\nu \geq 2\Delta$. With a GHV outlier model, $q_{FA}(10) \leq 10^{-3}$ as soon as $\nu \geq 6\sigma$. With the BGLV model, $\nu = 2\Delta$ ensures both $q_D(10,1) > 0.95$ and $q_{FA}(10) \leq 10^{-1}$.

In what follows, all results have been averaged over 2000 independent realizations of the WSN.

A. Performance of proposed DFD algorithm, ideal communication model

One considers here that $\varepsilon_I = \varepsilon_{II} = 0$, *i.e.*, channel access issues and packets losses are neglected.

First, the performance of a single round of the DFD algorithm described in Section III-A is evaluated. Figure 4 shows P_D as a function of P_{FA} parametrized in the threshold $\gamma \in (0, 1]$, for different values of the ratio $\bar{\mu}_d/\bar{\mu} \in \{0.02, 0.1\}$ and for the different outlier models presented in Section V-A. These curves are obtained with $\zeta = 1$. When γ increases, P_D and P_{FA} decrease in all cases, as shown by Lemma 4. For a given outlier model, the performance (in terms of detection accuracy) becomes worse as $\bar{\mu}_d/\bar{\mu}$ goes larger. Among the different outlier models, the GHV provides the worst performance. This comes from the fact that nodes with defective sensors may produce measurements that appear valid with a relatively high probability. Figure 4 also represents with straight lines the upper bounds for P_D and P_{FA} provided by Lemma 5 for all possible values of $\gamma \in (0, 1]$. For small values of γ , these bounds are relatively tight.

Figure 5 represents P_D and P_{FA} as a function of $\bar{\mu}$, with $\gamma \in \{0.8, 1\}$ and $\zeta = 1$, and for different values of $\bar{\mu}_d/\bar{\mu}$. Increasing $\bar{\mu}$ increases P_D , but also increases P_{FA} . For large values of $\bar{\mu}$, P_{FA} is unacceptable. Figure 6 represents a similar curve, but now, ζ is adapted for each value of $\bar{\mu}$ to have LODT involving less than 6 data in average. The main benefit is a decrease of P_{FA} when $\bar{\mu}$ increases. The price to be paid is a saturation of P_D when $\bar{\mu} > 6$.

The role played by the number of rounds on the achievable performance of the proposed DFD algorithms (single-decision and iterative) is then illustrated in Figures 7-9 corresponding to the outlier models UHV, GHV, and BGLV,

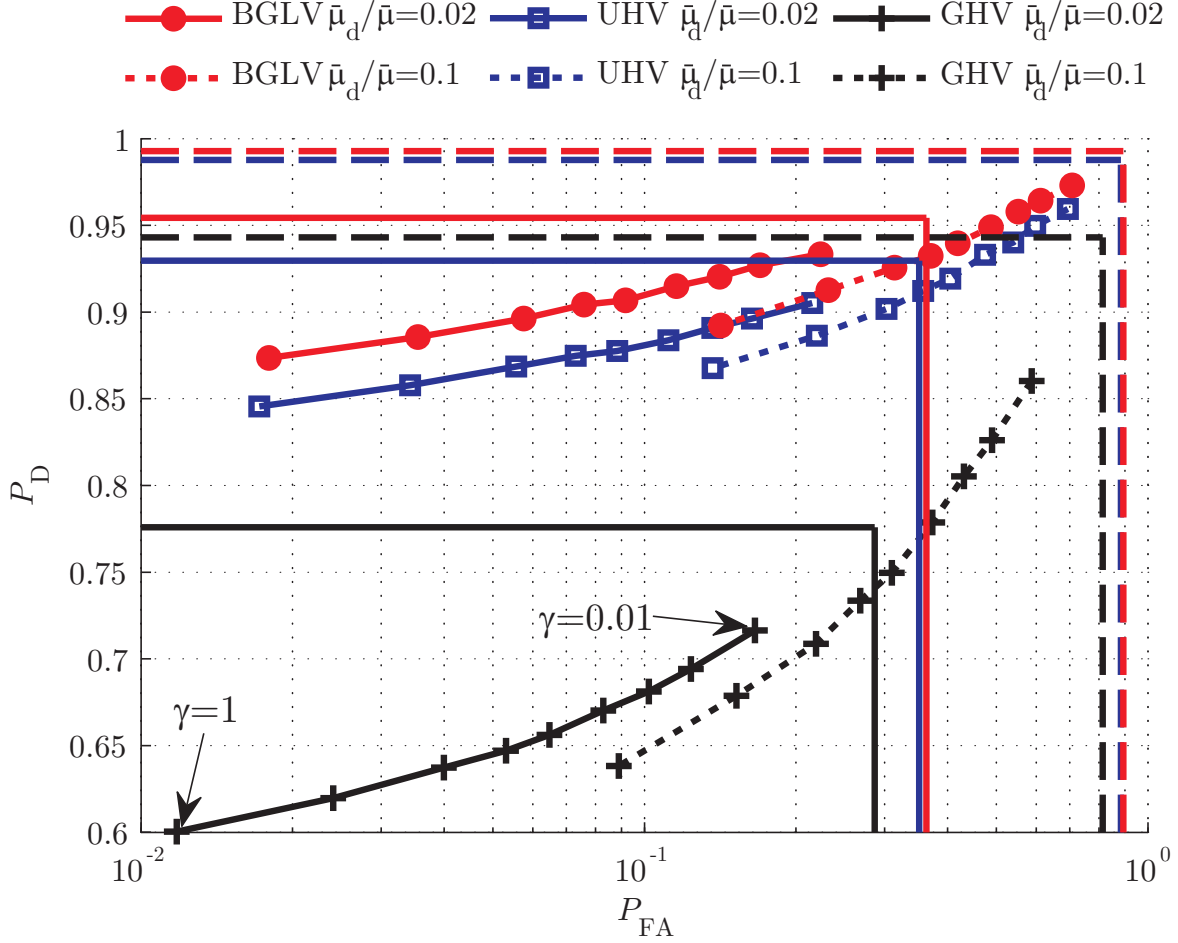


Fig. 4. P_D as a function of P_{FA} for a single round of the DFD of Section III-A parametrized in γ , for different $\bar{\mu}_d/\bar{\mu}$, for $\bar{\mu} = 6$; upper bounds for P_D and P_{FA} are also provided (straight lines)

respectively. P_D and P_{FA} are shown as functions of the round index $1 \leq \ell \leq 10$. One sets $\zeta = 1$ and $\gamma = 0.6$ in the iterative algorithm in all the cases, while the value of γ in the single-decision algorithm is adjusted so that both algorithms result to a similar P_D when $L = 10$. With this setting, one may evaluate the performance by comparing P_{FA} only. As can be observed, the single-decision DFD performs well when $\bar{\mu}_d/\bar{\mu}$ is small. However, when $\bar{\mu}_d/\bar{\mu} = 0.1$, P_{FA} is larger than 10% to keep a large P_D .

For the single decision algorithm, an oscillating behavior of P_D , depending on the iteration after which the decision is taken can be observed. This is due to the decision rule (19), where z_i/n_i , for small values of L , may vary significantly, since n_i and z_i are small. This effect vanishes as L increases, since n_i and z_i increase as well.

For the iterative algorithm, one considers $L_1 = 7$, and $L_2 = 3$. As can be observed, the seven iterative rounds of the first stage significantly reduces P_{FA} . The oscillations of P_D and P_{FA} are due to the fact that, during the iterative phase of the algorithm, a decreasing proportion of nodes in false alarm alternatively turns off and on in subsequent rounds until the algorithm switches to the single-decision phase. During the following three rounds, P_D improves rapidly, while P_{FA} remains small. This is due to the fact that most of the defective sensors have been detected and

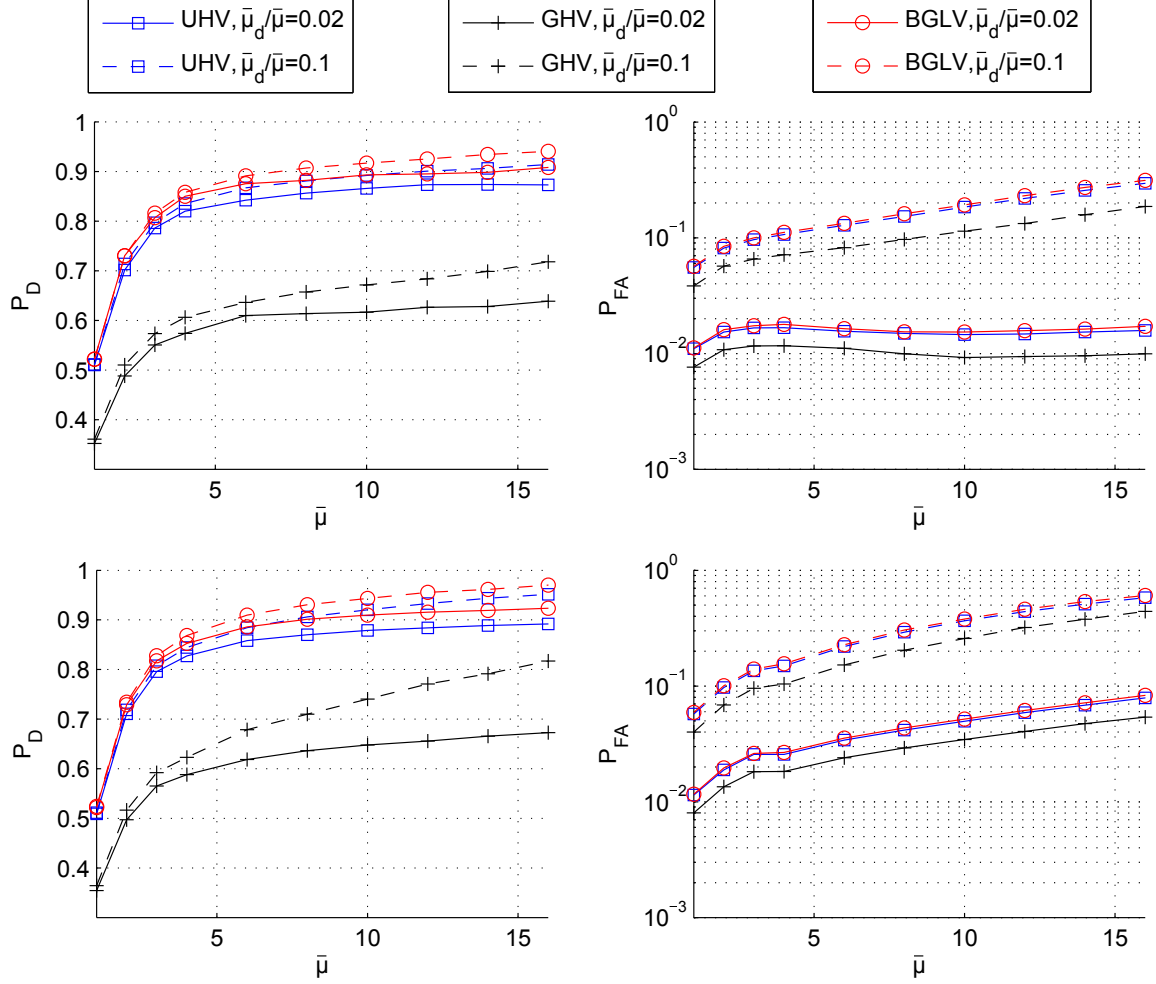


Fig. 5. P_D (left) and P_{FA} (right) as functions of $\bar{\mu}$ for a single round of DFD for different values of $\bar{\mu}_d/\bar{\mu}$ with $\gamma \in \{0.8, 1\}$ and $\zeta = 1$.

turned off during the previous iterative phase. Hence at the beginning of the single-decision rounds the residual $\bar{\mu}_d/\bar{\mu}$ is much lower than the initial $\bar{\mu}_d/\bar{\mu}$. Thus the iterative DFD algorithm performs better than the single-decision variant. For the same value of P_D , P_{FA} is almost one order of magnitude smaller with the iterative variant.

To better understand the oscillating behavior observed in Figures 7-9, one has to study the stability of the iterative DFD algorithm. Consider for example $\bar{\mu} = 6$ and the UHV outlier model with $\xi = 10$ and $\nu = 2\Delta$. Evaluating f_1 , f_2 , P_{FA} and P_D as a function of $\bar{\mu}_{00}$ and $\bar{\mu}_{10}$, the equilibrium can be characterized numerically. Consider $\zeta = 1$ and

$$\begin{cases} E_1(\mu_{00}, \mu_{10}) = \alpha_1 g_{FA}(\alpha_3 \mu_{00}, \alpha_3 \mu_{10}) - s_1(\alpha_3 \mu_{00}, \alpha_3 \mu_{10}), \\ E_2(\mu_{00}, \mu_{10}) = \alpha_1 g_D(\alpha_3 \mu_{00}, \alpha_3 \mu_{10}) - s_2(\alpha_3 \mu_{00}, \alpha_3 \mu_{10}), \end{cases}$$

deduced from (42-43), the equilibrium conditions may be rewritten as

$$\begin{cases} E_1(\mu_{00}^*, \mu_{10}^*) = 0, \\ E_2(\mu_{00}^*, \mu_{10}^*) = 0. \end{cases}$$

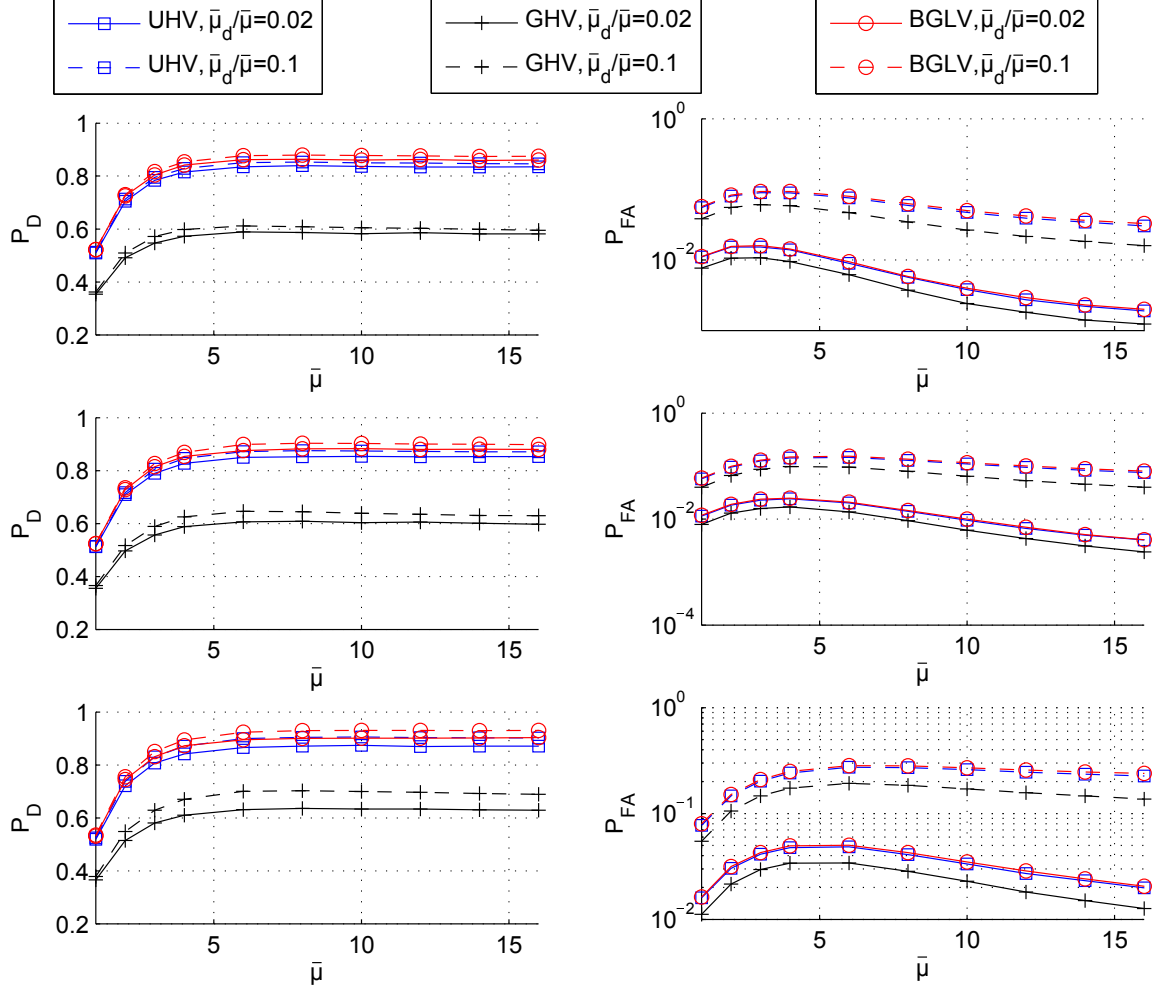


Fig. 6. P_D (left) and P_{FA} (right) as functions of $\bar{\mu}$ for a single round of DFD for different values of $\bar{\mu}_d/\bar{\mu}$ with $\gamma \in \{0.8, 1\}$ and ζ adapted to have LODT involving less than 6 neighboring data.

Figure 10 shows that $E_1(\mu_{00}^*, \mu_{10}^*) = E_2(\mu_{00}^*, \mu_{10}^*) = 0$ for the pairs (μ_{00}^*, μ_{10}^*) reported in Table III, with $\bar{\mu}_d/\bar{\mu} \in \{0.02, 0.1, 0.2\}$. The simulation results $(\mu_{00}^*/\bar{\mu}_g, \mu_{10}^*/\bar{\mu}_d)_s$ obtained from Figure 7 are also presented in Table III showing a very good match with the theoretical results $(\mu_{00}^*/\bar{\mu}_g, \mu_{10}^*/\bar{\mu}_d)_t$ obtained by solving (42-43).

The derivatives of f_1 , f_2 , P_{FA} and P_D with respect to μ_{00} and μ_{10} allow to get the matrix \mathbf{A} of the linearized model (51) and to evaluate its eigenvalues λ_1 and λ_2 . Table III shows that for three different values of the ratio $\bar{\mu}_d/\bar{\mu}$, the eigenvalues are within the unit circle. The linearized system is thus asymptotically stable. Moreover, the norms of the eigenvalues increase with $\bar{\mu}_d/\bar{\mu}$. The iterative DFD algorithm converges thus faster to equilibrium when $\bar{\mu}_d/\bar{\mu}$ is small.

B. Performance of proposed DFD algorithm, realistic communication model

In this section, channel access issues as well as packet losses are considered. For that purpose, we assume $\alpha_1 = 0.95$, $\varepsilon_1 = \varepsilon_{II} = 0.10$. This leads to $\alpha_2 \approx 0.9$, and $\alpha_3 \approx 0.85$. A node manages to access the channel in 95%

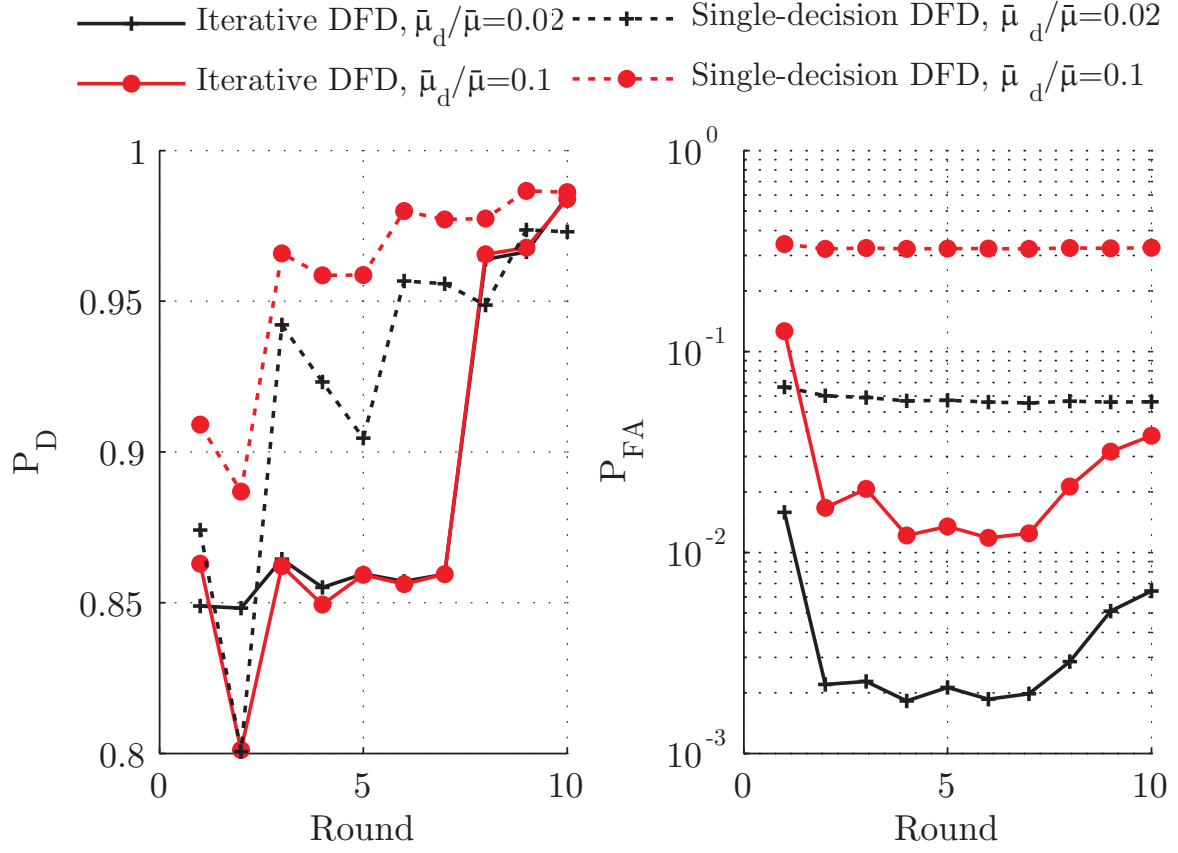


Fig. 7. P_D (left) and P_{FA} (right) as a function of the round index, with $\bar{\mu} = 6$, $\zeta = 1$, and $\bar{\mu}_d/\bar{\mu} \in \{0.02, 0.1\}$. The outlier model is UHV; For the single-decision algorithm, $\gamma = 0.65$, for the iterative algorithms, $\gamma = 0.6$, and $L_1 = 7$ and $L_2 = 3$.

TABLE III
ENTRIES OF THE MATRIX \mathbf{A} IN (51) AND OF ITS EIGENVALUES λ_1 AND λ_2 FOR THE UHV OUTLIER MODEL WHEN $\bar{\mu} = 6$, $\xi = 10$, AND $\nu = 2\Delta$

$\bar{\mu}_d/\bar{\mu}$	0.02	0.1
$\left(\frac{\mu_{00}^*}{\bar{\mu}_g}, \frac{\mu_{10}^*}{\bar{\mu}_d}\right)_t$	(0.998, 0.145)	(0.988, 0.145)
$\left(\frac{\mu_{00}^*}{\bar{\mu}_g}, \frac{\mu_{10}^*}{\bar{\mu}_d}\right)_s$	(0.995, 0.142)	(0.988, 0.142)
\mathbf{A}	$\begin{pmatrix} 0.011, & -2.02 \\ -0.001, & -0.006 \end{pmatrix}$	$\begin{pmatrix} 0.075 & -2.01 \\ -0.005 & -0.052 \end{pmatrix}$
(λ_1, λ_2)	(0.05, 0.046)	(0.13, 0.10)

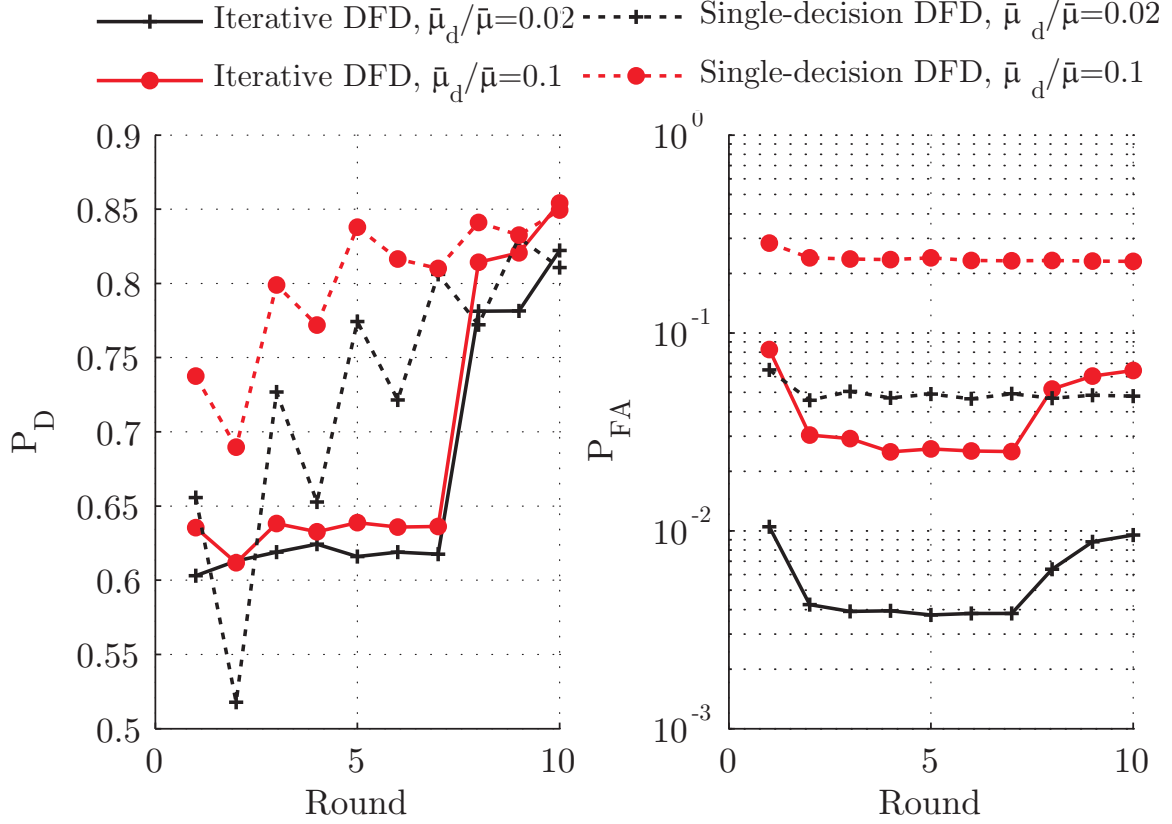


Fig. 8. P_D (left) and P_{FA} (right) as a function of the number of rounds, with $\bar{\mu} = 6$, $\zeta = 1$, and $\bar{\mu}_d/\bar{\mu} \in \{0.02, 0.1\}$. The outlier model is GHV. For the single-decision algorithm, $\gamma = 0.5$ as $\bar{\mu}_d/\bar{\mu} = 0.02$ and $\gamma = 0.59$ as $\bar{\mu}_d/\bar{\mu} = 0.1$. For the iterative algorithms, $\gamma = 0.6$, $L_1 = 7$, and $L_2 = 3$.

of the cases during a given phase of the considered algorithms. It has 90% chance to receive in Phase I a packet from one of its neighbors, and 85% chance to receive in Phase II a packet containing a decision involving the data it manages to broadcast in Phase I.

Figure 11 is the counterpart of Figure 7 for the UHV outlier model. P_D (left) and P_{FA} (right) are depicted as a function of the round index, with $\bar{\mu} = 6$, $\zeta = 1$, and $\bar{\mu}_d/\bar{\mu} \in \{0.02, 0.1\}$. For the single-decision algorithm, $\gamma = 0.65$, whereas for the iterative algorithms, $\gamma = 0.6$, and $L_1 = 7$ and $L_2 = 3$.

As far as the iterative algorithm is concerned, the values of P_D and P_{FA} are very close to those obtained with ideal communication conditions. During iterations, the oscillations are better damped. For example, when $\bar{\mu}_d/\bar{\mu} = 0.1$, it results $\lambda_1 = 0.034$ and $\lambda_2 = -0.117$.

For the single decision algorithm, performance are more degraded, *e.g.*, P_D is reduced from 0.97 to 0.91 when the more realistic communication model is considered.

C. Comparison with other DFD solutions

This section compares the proposed approach with two alternative DFD algorithms presented in [7] and in [16]. These reference DFD schemes have a LODT relatively close to that considered here. Nevertheless, a fair comparison

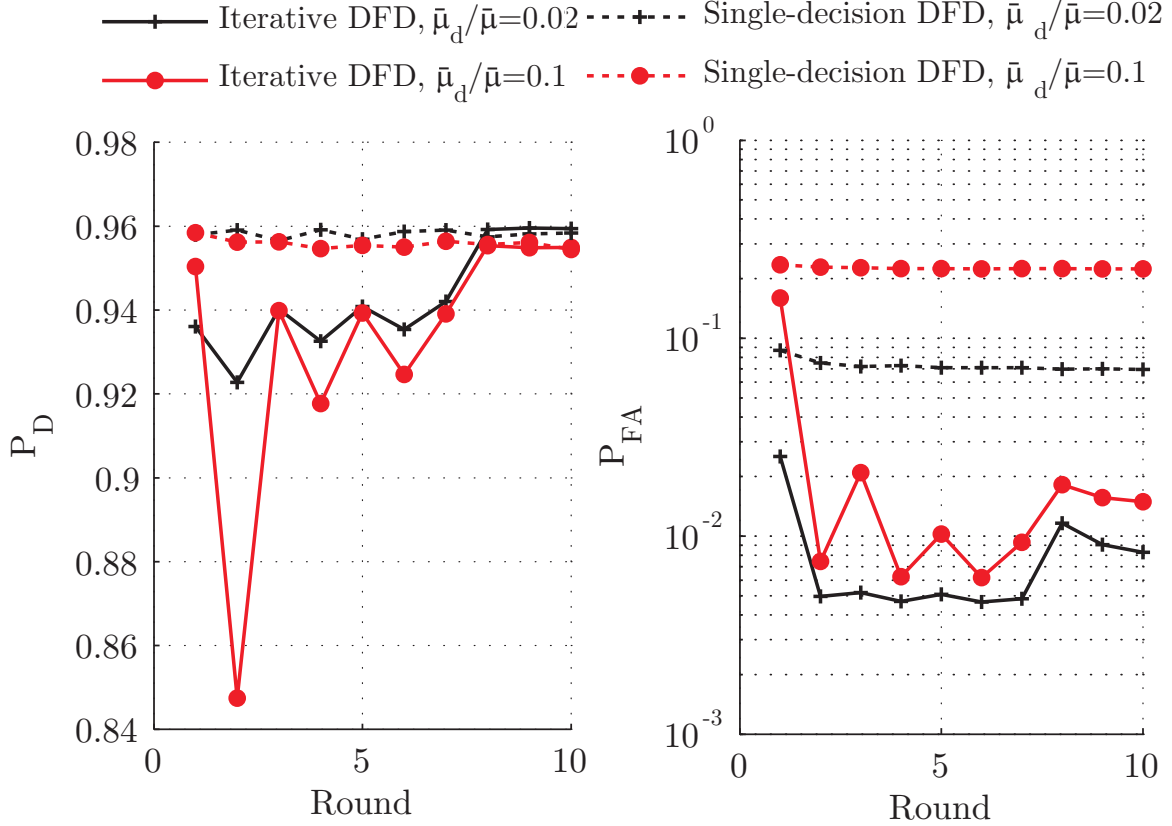


Fig. 9. P_D (left) and P_{FA} (right) as a function of the number of rounds, with $\bar{\mu} = 6$, $\zeta = 1$, and $\bar{\mu}_d/\bar{\mu} \in \{0.02, 0.1\}$. The outlier model is BGLV. For the single-decision algorithm, $\gamma = 0.68$ as $\bar{\mu}_d/\bar{\mu} = 0.02$ and $\gamma = 0.85$ as $\bar{\mu}_d/\bar{\mu} = 0.1$. For the iterative algorithms, $\gamma = 0.6$, $L_1 = 7$, and $L_2 = 3$.

between the two algorithms is quite difficult because the tests have some differences in their form and parameters.

First, ideal communication conditions are considered. In [7], in the first stage, each node exchanges its measurement with its neighbors and then performs a LODT. Compared to the proposed approach, Node i performs pairwise comparisons of its local measurement with all measurements received from its neighbors, to produce a binary value $c_{i,j}$ indicating whether there exists an outlier among m_i and m_j . Then, Node i decides whether its sensor is defective based on $c_{i,j}$, $j \in \mathcal{N}_i$. In a second stage, Node i broadcasts the estimate of its status, as well as its table of neighbors (including the index of Node j and $c_{i,j}$), which needs $(\bar{\mu} + 1)(1 + \log_2 |\mathcal{S}|)$ bits in average. This is of the same order of magnitude as the results indicated in Sections IV-B and IV-E.

Figure 12 compares the proposed DFD and the DFD algorithm in [7] in terms of P_D and P_{FA} , after 10 rounds of the algorithms, with $\bar{\mu} = 6$. In fact, the only parameter one can control in the DFD algorithm of [7] is the threshold of a local test (similar to ν in our case), whereas the performance of the proposed DFD depends on ν , γ , ζ , L_1 , and L_2 . For example, we fix ν (as the value in Table II), $L_1 = 7$, $L_2 = 3$ and choose different values of γ and $\zeta = 1$. As can be seen in Figures 12, the (P_D, P_{FA}) curve of the proposed DFD algorithm is in most of the cases above that of the DFD algorithm of [7].

Now packet losses and the channel access issues are taken into account. The DFD algorithm of [16] is considered,

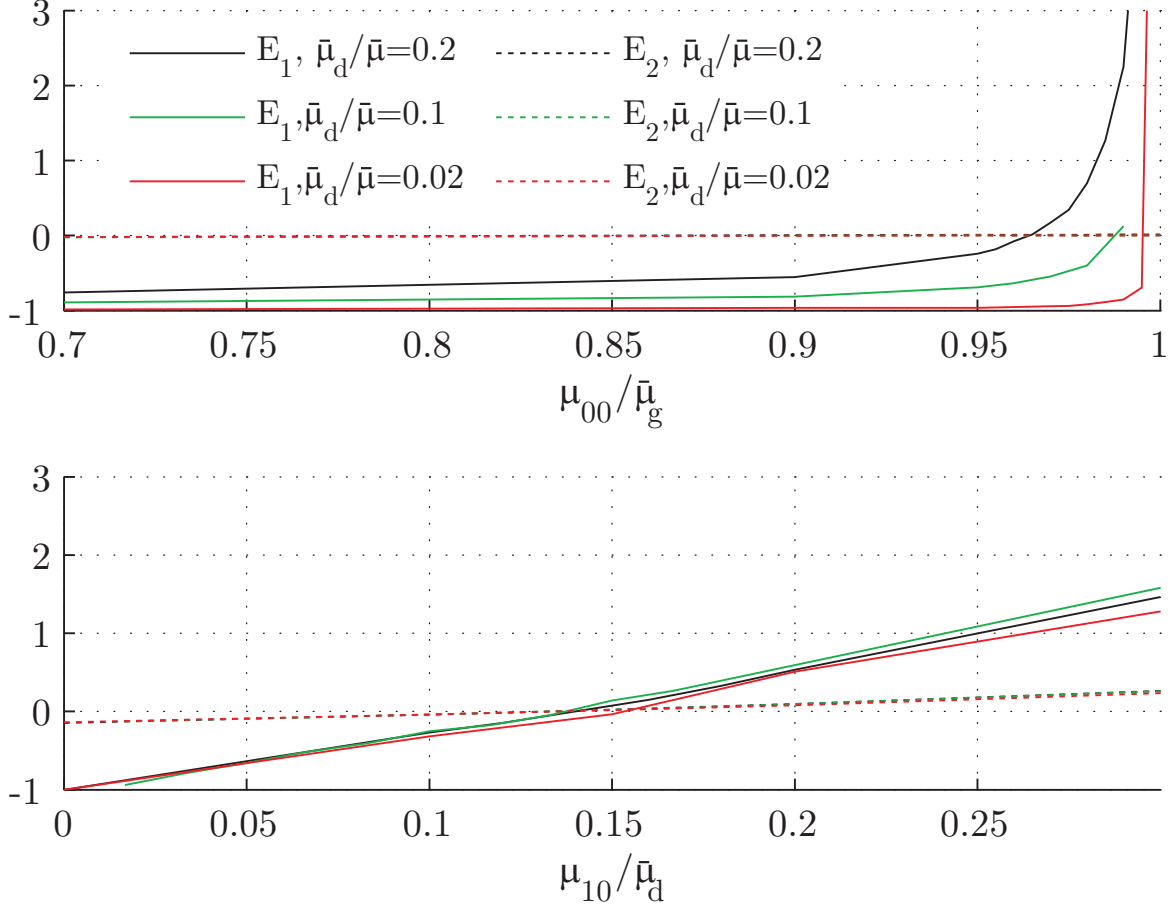


Fig. 10. Evolution of $E_1(\mu_{00}, \mu_{10}^*)$ and $E_2(\mu_{00}, \mu_{10}^*)$ as functions of $\mu_{00}/\bar{\mu}_g$ (top), and of $E_1(\mu_{00}^*, \mu_{10})$ and $E_2(\mu_{00}^*, \mu_{10})$ as functions of $\mu_{10}/\bar{\mu}_d$ (bottom), with $\bar{\mu}_d/\bar{\mu} \in \{0.02, 0.1, 0.2\}$ and $\zeta = 1$

since it is a modified version of that of [7] accounting for channel impairments. The UHV and BGLV models are considered for different values of the probability P_{success} of successful channel access and transmission. Results shown in Figure 13 compare also the performance of the proposed algorithm with and without transmission of nodes indices in $\mathcal{V}_i^{(\ell,1)}$. The performance of the algorithm in [16] is also indicated. With or without the knowledge of $\mathcal{V}_i^{(\ell,1)}$, the performance of the proposed algorithm are very close, especially when $P_{\text{success}} \geq 0.8$. The amount of data that needs to be transmitted between nodes can thus be reduced to one bit per node for the LODT result dissemination phase.

For the UHV model, the DFD algorithm in [16] provides lower P_{FA} for a similar P_{D} than the proposed algorithm only when P_{success} is close to 1. When P_{success} decreases the price to be paid in [16] for a low P_{FA} translates in a significantly degraded P_{D} . For the BGLV model, the DFD algorithm in [16] provides higher P_{FA} for a similar P_{D} for $P_{\text{success}} \in [0.6, 1]$ compared to the proposed approach. When $P_{\text{success}} < 0.6$, the value of P_{D} significantly decreases, even if P_{FA} decreases too. In this case, the proposed approach performs better for all values of P_{success} .

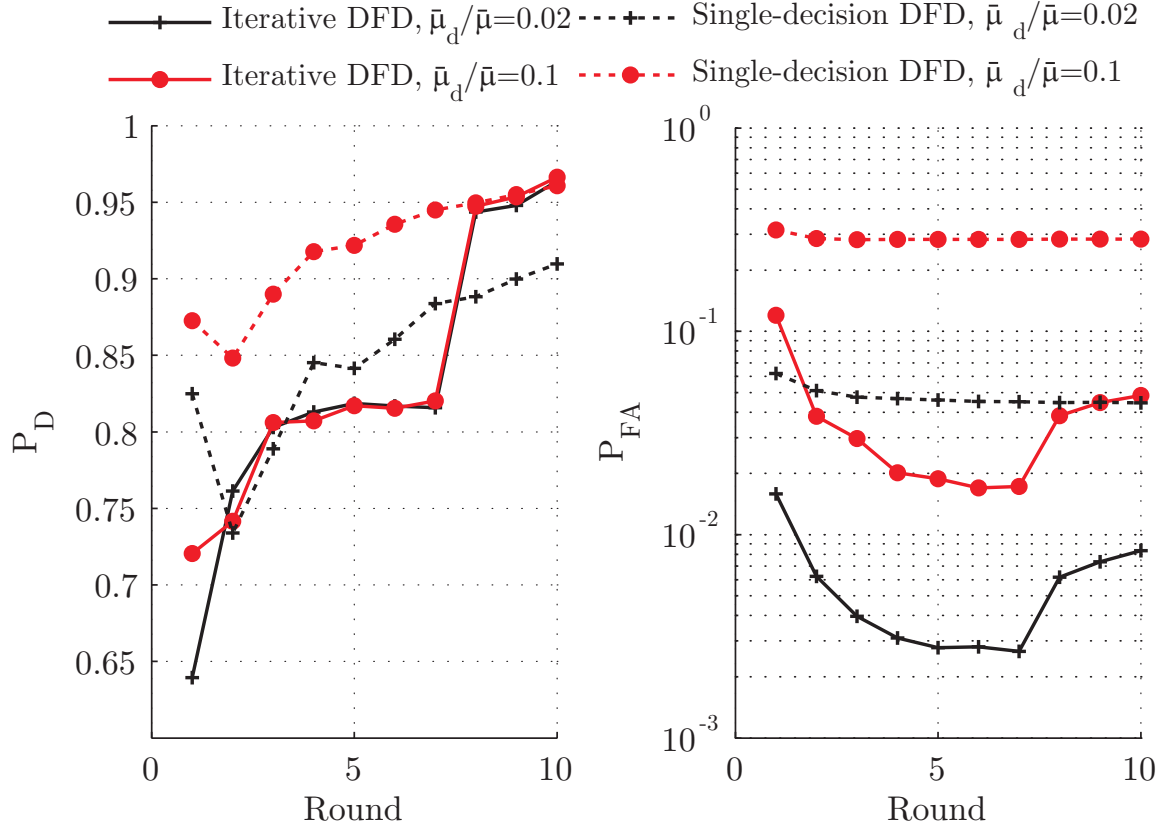


Fig. 11. P_D (left) and P_{FA} (right) as a function of the round index, with $\bar{\mu} = 6$, $\zeta = 1$, and $\bar{\mu}_d/\bar{\mu} \in \{0.02, 0.1\}$. The outlier model is UHV; For the single-decision algorithm, $\gamma = 0.65$, for the iterative algorithms, $\gamma = 0.6$, and $L_1 = 7$ and $L_2 = 3$.

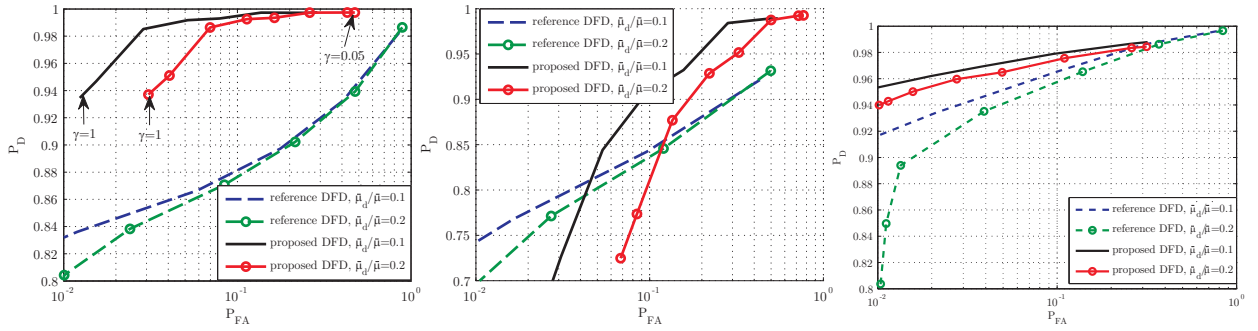


Fig. 12. P_D as a function of P_{FA} for the DFD algorithm in [7] and the proposed DFD after 10 rounds, with $\bar{\mu}_d/\bar{\mu} \in \{0.1, 0.2\}$ and the outlier models UHV (left), GHV (middle), and BGLV (right).

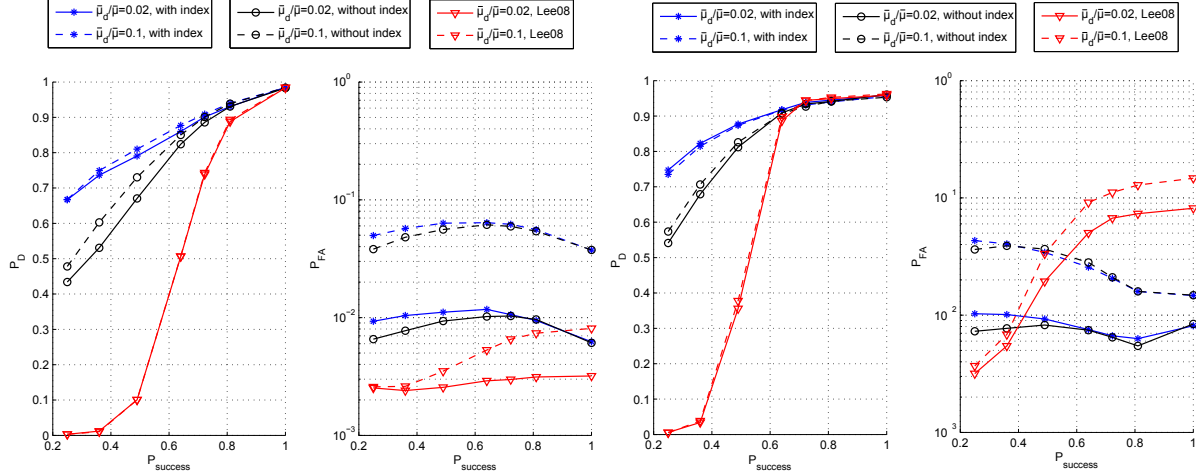


Fig. 13. Comparison of the iterative algorithms with or without the transmission of $\mathcal{V}_j^{(\ell,1)}$, as well as the DFD algorithm of [16], with $\bar{\mu} = 6$, $\zeta = 1$, and $\bar{\mu}_d/\bar{\mu} \in \{0.02, 0.1\}$. The outlier model is UHV (left) and BGLV (right). For the iterative algorithms, $\gamma = 0.6$, $L_1 = 7$, and $L_2 = 3$.

VII. CONCLUSIONS

This work proposed two variants of a two-stage DFD algorithm, which allows each node of a WSN to decide whether its sensor is producing outliers. The performance of the single-decision algorithm has been theoretically characterized. For the iterative variant, sufficient conditions to be satisfied by the LODT are identified to ensure existence of an equilibrium. These results are generic, since no specific form of the LODT needs to be considered. These conditions translate into upper bounds on $\bar{\mu}_g$ and $\bar{\mu}_d$. Conditions for the local asymptotic stability of this equilibrium are also provided. **The influence of channel access issues and packet losses has been analyzed.**

The DFD algorithm has been tested considering three outliers models and a simple local outlier detection algorithm. When the bounds on $\bar{\mu}_g$ and $\bar{\mu}_d$ are satisfied, the algorithms behave well. When they are not satisfied, some oscillating behavior appears, and an equilibrium is more difficult to reach. **Randomly selecting a subset of the received data to perform each LODT addresses this issue.**

The performance of both variants have been characterized by simulations, which enables to get some insights on the impact of the algorithm parameters (number of rounds, local test threshold, **probability of packet loss**) and of the network topology (density of faulty sensors, size of the neighborhood) on the trade-off between P_D and P_{FA} . **Comparisons with two alternative DFD algorithms have been performed, showing that for comparable or even lower traffic requirements, the proposed algorithm performs better in most of the cases.**

A way to account for the effect of node mobility during rounds is, *e.g.*, by an adaptation of ζ . Nevertheless, a precise analysis of the performance of both algorithm in case of significant node mobility is left for future research.

APPENDIX

A. Proof of Lemma 1

We start with the proof of (11).

Note that both q_D and q_{FA} represent the probability that the LODT yields 1. One has, for any $k \in \mathcal{A}$

$$\begin{aligned}
& \mathbb{P} \{T(\mathbf{M}_{\mathcal{A} \setminus \{k\}}) = 1 \mid \varphi(\mathcal{A}) = 0\} \\
&= \mathbb{P} \{T(\mathbf{M}_{\mathcal{A} \setminus \{k\}}) = 1, T(\mathbf{M}_{\mathcal{A}}) = 0 \mid \varphi(\mathcal{A}) = 0\} \\
&\quad + \mathbb{P} \{T(\mathbf{M}_{\mathcal{A} \setminus \{k\}}) = 1, T(\mathbf{M}_{\mathcal{A}}) = 1 \mid \varphi(\mathcal{A}) = 0\} \\
&\stackrel{(a)}{=} \mathbb{P} \{T(\mathbf{M}_{\mathcal{A} \setminus \{k\}}) = 1, T(\mathbf{M}_{\mathcal{A}}) = 1 \mid \varphi(\mathcal{A}) = 0\} \\
&\leq \mathbb{P} \{T(\mathbf{M}_{\mathcal{A}}) = 1 \mid \varphi(\mathcal{A}) = 0\},
\end{aligned} \tag{57}$$

where (a) is obtained from Property 2. Combining (7) and (57), one gets

$$q_{FA}(\mathbf{M}_{\mathcal{A} \setminus \{K\}}) \leq q_{FA}(\mathbf{M}_{\mathcal{A}}). \tag{58}$$

Suppose that $|\mathcal{A}| = n_g + 1$ with $n_g \geq 1$, then based on Property 1, (58) is equivalent to $q_{FA}(n_g) \leq q_{FA}(n_g + 1)$, which proves (11).

One may prove (12) and (13) in a similar way, by considering $\varphi(\mathcal{A}) = 1$ in the derivations.

B. Analysis of the decision rule (19)

At the end of Algorithm 1, each node estimates its status from the results of the LODTs, which may be gathered in the vector

$$\mathbf{Y}_i^{(L)} = \left[Y_j^{(\ell)} \right]_{j \in \mathcal{B}_i^{(\ell)}, 1 \leq \ell \leq L}. \tag{59}$$

Let $\mathbf{y}_i^{(L)}$ be one realization of $\mathbf{Y}_i^{(L)}$. In what follows, one will show that the decision rule (19) corresponds, under some simplifying assumptions, to

$$\hat{\theta}_i = \arg \max_{\phi \in \{0,1\}} c_\phi \mathbb{P} \left\{ \mathbf{Y}_i^{(L)} = \mathbf{y}_i^{(L)} \mid \theta_i = \phi \right\}, \tag{60}$$

where c_ϕ is some weight. If $c_0 = c_1 = 1$, then (60) is a maximum likelihood estimate of θ_i from $\mathbf{y}_i^{(L)}$; if $c_\phi = \mathbb{P} \{ \theta_i = \phi \}$, then (60) is the maximum *a posteriori* estimate of θ_i .

Expressing the likelihoods in (60) is complicated due to the correlation between the components of $\mathbf{Y}_j^{(L)}$. Nevertheless, assuming that the $Y_j^{(\ell)}$ s are iid, one obtains

$$\mathbb{P} \left\{ \mathbf{Y}_i^{(L)} = \mathbf{y}_i^{(L)} \mid \theta_i = \phi \right\} = \prod_{\ell=1}^L \prod_{j \in \mathcal{B}_i^{(\ell)}} \mathbb{P} \left\{ Y_j^{(\ell)} = y_j^{(\ell)} \mid \theta_i = \phi \right\}. \tag{61}$$

Assuming further that $\mathbb{P} \left\{ Y_j^{(\ell)} = y_j^{(\ell)} \mid \theta_i = \phi \right\} = \mathbb{P} \left\{ Y_i^{(\ell)} = y_i^{(\ell)} \mid \theta_i = \phi \right\}$ for all $j \in \mathcal{B}_i^{(\ell)}$, which can be justified by Property 1 when the number of measurements considered by Node j to get $Y_j^{(\ell)}$ is equal to the number of measurements by Node i to get $Y_i^{(\ell)}$. Then for all $j \in \mathcal{B}_i^{(\ell)}$, one gets

$$\mathbb{P} \left\{ Y_j^{(\ell)} = 1 \mid \theta_i = 0 \right\} = f_1(\bar{\mu}_g, \bar{\mu}_d), \tag{62}$$

$$\mathbb{P} \left\{ Y_j^{(\ell)} = 1 \mid \theta_i = 1 \right\} = f_2(\bar{\mu}_g, \bar{\mu}_d). \tag{63}$$

Let $z_i = \sum_{\ell=1}^L \sum_{j \in \mathcal{B}_i^{(\ell)}} y_j^{(\ell)} = z_i$ and $n_i = \sum_{\ell=1}^L |\mathcal{B}_i^{(\ell)}|$, then from (61-62) one obtains

$$\frac{c_0 \mathbb{P} \left\{ \mathbf{Y}_i^{(L)} = \mathbf{y}_i^{(L)} \mid \theta_i = 0 \right\}}{c_1 \mathbb{P} \left\{ \mathbf{Y}_i^{(L)} = \mathbf{y}_i^{(L)} \mid \theta_i = 1 \right\}} = \frac{c_0 f_1^{z_i} (1 - f_1)^{n_i - z_i}}{c_1 f_2^{z_i} (1 - f_2)^{n_i - z_i}}. \quad (64)$$

Therefore, one obtains the following decision rule

$$\hat{\theta}_i = \begin{cases} 0, & \text{if } \frac{z_i}{n_i} < \frac{\frac{1}{n_i} \log \frac{c_1}{c_0} - \log \frac{1-f_1}{1-f_2}}{\log \frac{f_1(1-f_2)}{f_2(1-f_1)}}, \\ 1, & \text{otherwise.} \end{cases} \quad (65)$$

When $c_0 = c_1 = 1$, the threshold corresponding to the maximum likelihood estimate is

$$\gamma_{\text{ML}} = \frac{\log \frac{1-f_1}{1-f_2}}{\log \frac{1-f_1}{1-f_2} + \log \frac{f_2}{f_1}}. \quad (66)$$

When $c_\phi = \mathbb{P} \{ \theta_i = \phi \}$, the threshold corresponding to the maximum *a posteriori* estimate of θ depends on n_i as follows

$$\gamma_{\text{MAP}}(n_i) = \frac{\log \frac{1-f_1}{1-f_2} - \frac{1}{n_i} \log \frac{\bar{\mu}_d}{\bar{\mu}_g}}{\log \frac{1-f_1}{1-f_2} + \log \frac{f_2}{f_1}}. \quad (67)$$

One notes that $\lim_{n_i \rightarrow \infty} \gamma_{\text{MAP}}(n_i) = \gamma_{\text{ML}}$.

Figure 14 represents the theoretical values of γ_{ML} , $\gamma_{\text{MAP}}(\bar{\mu})$ with $L = 1$, and $\gamma_{\text{MAP}}(\bar{\mu})$ with $l = 10$ as functions of $\bar{\mu}$ for different values of $\bar{\mu}_d/\bar{\mu}_g$. For a constant value of $\bar{\mu}_g$, γ increases with $\bar{\mu}_d$. Moreover, γ_{ML} and $\gamma_{\text{MAP}}(\bar{\mu})$ with $L = 10$ behave very similarly, increasing with $\bar{\mu}$. Finally, $\gamma_{\text{MAP}}(\bar{\mu})$ with $L = 1$ is always larger than the other thresholds: when a limited number of measurements is available, this limits the probability of false alarm.

C. Multi-hop Algorithm

When the WSN is sparse, *i.e.*, when $\bar{\mu}$ is small, the performance of DFD algorithms may be poor, due to a reduced amount of data to perform LODT. This appendix presents the modifications to be performed in the DFD algorithm to allow multi-hop data collection and LODT result dissemination. It is a variant of multi-hop dissemination protocol, see, *e.g.*, [2, Chap. 4.5], where data aggregation in packets is performed to limit the number of transmitted packets.

This feature requires that each node acts as a router to forward the packets it has received. This forwarding has to be limited using some time-to-live information T^{TL} associated to each data. T^{TL} has to be initialized to H , the maximum number of allowed hops. Let $\mathbf{p}_i^{(\ell, k)}$ be the payload of the k -th packet transmitted by Node i , during Phase I in the ℓ -th round. Initially $k = 1$ and one has

$$\mathbf{p}_i^{(\ell, 1, 1)} = \left\{ \left(T_i^{\text{TL}}, \quad m_i^{(\ell)} \right) \right\}, \quad (68)$$

with $T_i^{\text{TL}} = H$. The identification of Node i , stored in Addr_i , is included in the packet header. Node i has received packets from its neighbors in the set $\mathcal{N}_i^{(\ell, 1, 1)}$ during the first transmission of Phase I of Round ℓ . The corresponding data and associated T^{TL} are denoted $\left(T_j^{\text{TL}}, m_j^{(\ell)} \right)$, $j \in \mathcal{N}_i^{(\ell, 1, 1)} \subset \mathcal{S}$. Node i then needs to forward all data which have a $T_j^{\text{TL}} > 0$. To save bandwidth resources, these data are aggregated into one packet to get

$$\mathbf{p}_i^{(\ell, 1, 2)} = \bigcup_{j \in \mathcal{N}_i^{(\ell, 1, 1)} \text{ st } T_j^{\text{TL}} > 0} \left\{ \left(\text{Addr}_j, \quad T_j^{\text{TL}} - 1, \quad m_j^{(\ell)} \right) \right\}. \quad (69)$$

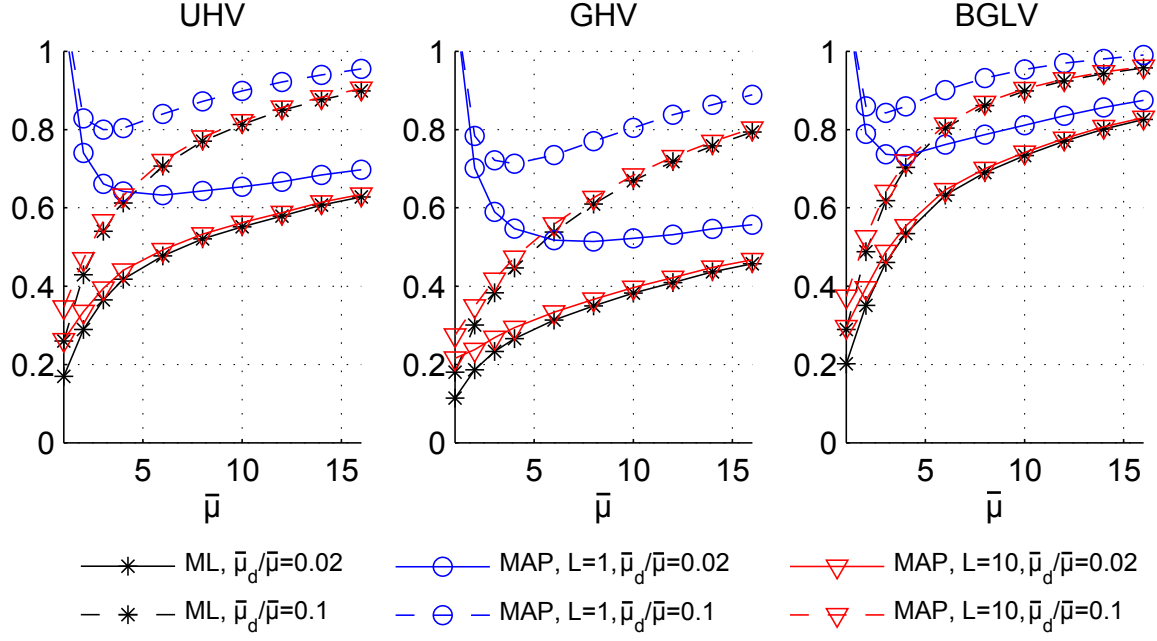


Fig. 14. γ_{ML} , $\gamma_{\text{MAP}}(\bar{\mu})$ with $L = 1$, and $\gamma_{\text{MAP}}(\bar{\mu})$ with $L = 10$ as functions of $\bar{\mu}$ for different values of $\bar{\mu}_d/\bar{\mu}$.

During k -th transmission with $k \geq 2$, packets are transmitted similarly, accounting for newly received data.

A similar process is performed during Phase II where LODT results are broadcast with limited hop count. Now, the payload of the first packet transmitted by Node i , during Phase II in the ℓ -th round is

$$\mathbf{p}_i^{(\ell, \text{II}, 1)} = \left[T_i^{\text{TL}}, y_i^{(\ell)}, |\mathcal{V}_i^{(\ell, 1)}|, \bigcup_{j \in \mathcal{V}_i^{(\ell, 1)}} \{\text{Addr}_j\} \right]. \quad (70)$$

In (70), T_i^{TL} is again initialized at H . The number of nodes participating to the LODT results $y_i^{(\ell)}$ as well as their addresses are also indicated. For the k -th transmitted packets, payload containing LODT outcomes with $T^{\text{TL}} > 0$ can again be aggregated as done in (69).

D. Proof of (44)

Consider a bounded sequence $0 \leq B(n) \leq 1$ for any $n \in \mathbb{N}$ and a real-valued function $G_N(\mu) = \sum_{n=0}^N B(n) \cdot \frac{\mu^n}{n!} \exp(-\mu)$ with $\mu \geq 0$. First, one evaluates the derivative of $G_N(\mu)$,

$$\begin{aligned} \frac{dG_N(\mu)}{d\mu} &= \exp(-\mu) \cdot \\ &\left(\sum_{n=0}^N (B(n+1) - B(n)) \frac{\mu^n}{n!} - B(N+1) \frac{\mu^N}{N!} \right). \end{aligned} \quad (71)$$

Second, one shows that $\frac{dG_N(\mu)}{d\mu}$ converges uniformly to

$$H(\mu) = \sum_{n=0}^{\infty} (B(n+1) - B(n)) \frac{\mu^n}{n!} \exp(-\mu). \quad (72)$$

Let $\varepsilon \in \mathbb{R}^+$ be an arbitrary number, then for any $\mu \geq 0$, there exists $N_\mu \in \mathbb{N}^+$, such that

$$\sum_{n=N_\mu}^{\infty} \frac{\mu^n}{n!} \exp(-\mu) < \varepsilon. \quad (73)$$

One has

$$\begin{aligned} \left| \frac{dG_{N_\mu}(\mu)}{d\mu} - H(\mu) \right| &\leq B(N_\mu + 1) \cdot \frac{\mu^{N_\mu}}{N_\mu!} \exp(-\mu) + \\ &\quad \sum_{n=N_\mu+1}^{\infty} |B(n+1) - B(n)| \cdot \frac{\mu^n}{n!} \exp(-\mu) \\ &\stackrel{(a)}{\leq} \sum_{n=N_\mu}^{\infty} \frac{\mu^n}{n!} \exp(-\mu) \stackrel{(b)}{<} \varepsilon, \end{aligned}$$

where (a) comes from the fact that $0 \leq B(n) \leq 1$ and $-1 \leq B(n+1) - B(n) \leq 1$ for any $n \in \mathbb{N}$, and (b) is by (73). Similarly, one can also show that $G_N(\mu)$ converges uniformly to $G(\mu) = \sum_{n=0}^{\infty} B(n) \cdot \frac{\mu^n}{n!} \exp(-\mu)$ for any $\mu \geq 0$.

Therefore, according to [25, Thm 7.17], one obtains that $\lim_{N \rightarrow \infty} \frac{dG_N(\mu)}{d\mu} = H(\mu)$, $G(\mu)$ is differentiable for any $\mu \geq 0$ and $\frac{dG(\mu)}{d\mu} = H(\mu)$.

E. Proof of Lemma 5

Let $\mathcal{R}_i = \{j \in \mathcal{S} \text{ such that } r_{i,j} \leq 2R_0\}$ be the set of nodes at a distance to the reference Node i less than $2R_0$, including i itself. For any $j \in \mathcal{B}_i$, $\mathcal{V}_j^{(1)} \subseteq \mathcal{R}_i$. Let $\mathbf{M}_{\mathcal{R}_i}$ and $\mathbf{M}_{\mathcal{V}_j^{(1)}}$ be the vectors of data provided by the sensors in \mathcal{R}_i and $\mathcal{V}_j^{(1)}$ respectively. From Property 2, one has for all \mathcal{U}_j , $j \in \mathcal{U}_i$

$$\mathbb{P}\left\{T\left(\mathbf{M}_{\mathcal{V}_j^{(1)}}\right) = 0 \mid T\left(\mathbf{M}_{\mathcal{R}_i}\right) = 0\right\} = 1, \quad \forall j \in \mathcal{B}_i. \quad (74)$$

To lighten notations, define $Z_i = \sum_{j \in \mathcal{B}_i} T\left(\mathbf{M}_{\mathcal{V}_j^{(1)}}\right)$. From (74), one deduces

$$\mathbb{P}\{Z_i = 0 \mid T(\mathbf{M}_{\mathcal{R}_i}) = 0\} = 1. \quad (75)$$

Then

$$\begin{aligned} \mathbb{P}\{Z_i = 0\} &= \sum_{y=0}^1 \mathbb{P}\{Z_i = 0, T(\mathbf{M}_{\mathcal{R}_i}) = y\} \\ &\geq \mathbb{P}\{Z_i = 0, T(\mathbf{M}_{\mathcal{R}_i}) = 0\} \\ &= \mathbb{P}\{Z_i = 0 \mid T(\mathbf{M}_{\mathcal{R}_i}) = 0\} \cdot \mathbb{P}\{T(\mathbf{M}_{\mathcal{R}_i}) = 0\} \\ &= \mathbb{P}\{T(\mathbf{M}_{\mathcal{R}_i}) = 0\}. \end{aligned} \quad (76)$$

On the other hand

$$\mathbb{P}\{Z_i < (N_g + N_d) \gamma\} \geq \mathbb{P}\{Z_i = 0\}. \quad (77)$$

Combining (76) and (77), one gets

$$\begin{aligned} \mathbb{P}\{Z_i \geq (N_g + N_d) \gamma\} &\leq 1 - \mathbb{P}\{Z_i = 0\} \\ &\leq \mathbb{P}\{T(\mathbf{M}_{\mathcal{R}_i}) = 1\}, \end{aligned} \quad (78)$$

independently of the status of Node i . When $\theta_i = 0$,

$$\begin{aligned} P_{\text{FA}} \left(\mu_{\text{g}}^{(\text{II})}, \mu_{\text{d}}^{(\text{II})} \right) &\leq \mathbb{P} \{ T(\mathbf{M}_{\mathcal{R}_i}) = 1 \mid \theta_i = 0 \} \\ &= \sum_{n_{\text{g}}=1}^{\infty} \sum_{n_{\text{d}}=0}^{\infty} h(n_{\text{g}}, n_{\text{d}}) \frac{(4\bar{\mu}_{\text{d}})^{n_{\text{d}}-1} (4\bar{\mu}_{\text{g}})^{n_{\text{g}}}}{(n_{\text{g}}-1)!n_{\text{d}}!} \exp(-4\bar{\mu}_{\text{d}} - 4\bar{\mu}_{\text{g}}), \end{aligned}$$

leading to (33). Note that the area of the disk associated to \mathcal{R}_i is $4\pi R_0^2$, thus in average $|\mathcal{R}_i \cap \mathcal{G}|$ is $4\bar{\mu}_{\text{g}}$ and in average $|\mathcal{R}_i \cap \mathcal{D}|$ is $4\bar{\mu}_{\text{d}}$. One may show (34) in a similar way.

F. Proof of Lemma 8

To prove that $g_{\text{FA}}(\mu_{00}, \mu_{10})$ as defined in (47) is monotone increasing in μ_{00} , one has to show that

$$\begin{aligned} \frac{\partial g_{\text{FA}}}{\partial \mu_{00}} &= \frac{\mu_{00}}{\alpha_3 \zeta \bar{\mu}_{\text{g}} - \mu_{00}} \sum_{n_{\text{g}}=1}^{\infty} \sum_{n_{\text{d}}=0}^{\infty} \frac{\mu_{10}^{n_{\text{d}}} \mu_{00}^{n_{\text{g}}-1} \exp(-\mu_{10} - \mu_{00})}{n_{\text{d}}! (n_{\text{g}}-1)!} \\ &\quad \left(\tau_{\text{FA}}(n_{\text{g}}+1, n_{\text{d}}) + \frac{\mu_{00}^2 - \alpha_3 \zeta \mu_{00} \bar{\mu}_{\text{g}} + \alpha_3 \zeta \bar{\mu}_{\text{g}}}{\mu_{00} (\alpha_3 \zeta \bar{\mu}_{\text{g}} - \mu_{00})} \tau_{\text{FA}}(n_{\text{g}}, n_{\text{d}}) \right) \end{aligned} \quad (79)$$

is strictly positive. Therefore, a sufficient condition to have $\frac{\partial g_{\text{FA}}}{\partial \mu_{00}} > 0$ is

$$\frac{\tau_{\text{FA}}(n_{\text{g}}+1, n_{\text{d}})}{\tau_{\text{FA}}(n_{\text{g}}, n_{\text{d}})} > -\frac{\mu_{00}^2 - \alpha_3 \zeta \mu_{00} \bar{\mu}_{\text{g}} + \alpha_3 \zeta \bar{\mu}_{\text{g}}}{(\alpha_3 \zeta \bar{\mu}_{\text{g}} - \mu_{00}) \mu_{00}}. \quad (80)$$

Since

$$-\frac{\mu_{00}^2 - \alpha_3 \zeta \mu_{00} \bar{\mu}_{\text{g}} + \alpha_3 \zeta \bar{\mu}_{\text{g}}}{(\alpha_3 \zeta \bar{\mu}_{\text{g}} - \mu_{00}) \mu_{00}} = 1 - \frac{\alpha_3 \zeta \bar{\mu}_{\text{g}}}{(\alpha_3 \zeta \bar{\mu}_{\text{g}} - \mu_{00}) \mu_{00}} \leq 1 - \frac{4}{\alpha_3 \zeta \bar{\mu}_{\text{g}}},$$

a sufficient condition to have (80) is

$$\frac{\tau_{\text{FA}}(n_{\text{g}}+1, n_{\text{d}})}{\tau_{\text{FA}}(n_{\text{g}}, n_{\text{d}})} > \left(1 - \frac{4}{\alpha_3 \zeta \bar{\mu}_{\text{g}}} \right)$$

for any $n_{\text{g}} \geq 1$ and $n_{\text{d}} \geq 0$, which is equivalent to

$$\min_{n_{\text{g}} \geq 1, n_{\text{d}} \geq 0} \frac{\tau_{\text{FA}}(n_{\text{g}}+1, n_{\text{d}})}{\tau_{\text{FA}}(n_{\text{g}}, n_{\text{d}})} > 1 - \frac{4}{\alpha_3 \zeta \bar{\mu}_{\text{g}}}. \quad (81)$$

One needs to find a lower bound of $\tau_{\text{FA}}(n_g + 1, n_d)$ as a function of $\tau_{\text{FA}}(n_g, n_d)$. Introducing $\mathcal{I} = \{\theta_i = 0, N_g = n_g + 1, N_d = n_d\}$, by definition,

$$\begin{aligned}
\tau_{\text{FA}}(n_g + 1, n_d) &= \mathbb{P} \left\{ \frac{\sum_{j \in \mathcal{B}_i} Y_j}{n_g + n_d + 1} = 1 \middle| \mathcal{I} \right\} \\
&\stackrel{(a)}{=} \mathbb{P} \left\{ \frac{\sum_{j \in \mathcal{B}_i \setminus \{K\}} Y_j}{n_g + n_d} = 1 \middle| \theta_K = 0, \mathcal{I} \right\} \\
&\quad \cdot \mathbb{P} \left\{ Y_K = 1 \middle| \frac{\sum_{j \in \mathcal{B}_i \setminus \{K\}} Y_j}{n_g + n_d} = 1, \theta_K = 0, \mathcal{I} \right\} \\
&\stackrel{(b)}{\geq} \mathbb{P} \left\{ \frac{\sum_{j \in \mathcal{B}_i \setminus \{K\}} Y'_j}{n_g + n_d} = 1 \middle| \theta_K = 0, \mathcal{I} \right\} \\
&\quad \cdot \mathbb{P} \left\{ Y_K = 1 \middle| \frac{\sum_{j \in \mathcal{B}_i \setminus \{K\}} Y_j}{n_g + n_d} = 1, \theta_K = 0, \mathcal{I} \right\} \\
&\stackrel{(c)}{=} \mathbb{P} \left\{ \frac{\sum_{j \in \mathcal{B}_i \setminus \{K\}} Y'_j}{n_g + n_d} = 1 \middle| \theta_i = \theta_K = 0, N'_g = n_g, N_d = n_d \right\} \\
&\quad \cdot \mathbb{P} \left\{ Y_K = 1 \middle| \frac{\sum_{j \in \mathcal{B}_i \setminus \{K\}} Y_j}{n_g + n_d} = 1, \theta_K = 0, \mathcal{I} \right\} \\
&\stackrel{(d)}{=} \tau_{\text{FA}}(n_g, n_d) \cdot \mathbb{P} \left\{ Y_K = 1 \middle| \frac{\sum_{j \in \mathcal{B}_i \setminus \{K\}} Y_j}{n_g + n_d} = 1, \theta_K = 0, \mathcal{I} \right\} \\
&\stackrel{(e)}{\geq} \tau_{\text{FA}}(n_g, n_d) \cdot \mathbb{P} \left\{ Y_K = 1 \middle| Y_i = 1, \theta_K = 0, \mathcal{I} \right\},
\end{aligned} \tag{82}$$

where the considered \mathcal{B}_i is such that $\theta_i = 0$, $N_g = n_g + 1$, and $N_d = n_d$. For all $j \in \mathcal{B}_i$, the LODTs are performed based on the data vector $\mathbf{M}_{\mathcal{V}_j^{(i)}}$ with outcome Y_j . In (a), $K \neq i$ is a random node in $\mathcal{N}_i \cap \mathcal{G}$; such node exists since $n_g \geq 1$. In (b), $Y'_j = T(\mathbf{M}_{\mathcal{V}_j^{(i)} \setminus \{K\}})$ and one uses the results of Lemma 1 to get $\mathbb{P}\{Y_j = 1\} \geq \mathbb{P}\{Y'_j = 1\}$ for all $j \in \mathcal{B}_i \setminus \{K\}$. To get (c), one uses the fact that Node K , equipped with a good sensor, is not used to get Y'_j and (82) accounts only for the presence of $N'_g = n_g$ nodes with good sensors. Then (d) is by definition of $\tau_{\text{FA}}(n_g, n_d)$ and see Appendix G for the proof of (e). According to (81) and (83), the first statement of Lemma 8 is proved.

In the similar way, a sufficient condition to have $\mu_{10} P_D(\mu_{00}, \mu_{10}) / (\bar{\mu}_d - \mu_{10})$ an increasing function of μ_{10} is that

$$\begin{aligned}
\frac{\tau_D(n_g, n_d + 1)}{\tau_D(n_g, n_d)} &\geq \mathbb{P}\{Y_k = 1 \mid \theta_i = \theta_k = 1\} \\
&\geq q_D(0, 2) > 1 - \frac{4}{\alpha_3 \zeta \bar{\mu}_d},
\end{aligned} \tag{84}$$

which corresponds to the second statement of Lemma 8.

G. Proof of $\mathbb{P}\{Y_{j_1} = 1\} \leq \mathbb{P}\{Y_{j_1} = 1 \mid Y_i = 1, Y_{j_2} = 1, \dots\}$

In this section, one aims to prove the following lemma.

Lemma 11. Consider a LODT satisfying Properties 1 and 2, and some Node i . For any set of distinct indexes $\{i, j_1, j_2, \dots\} \subset \mathcal{B}_i$, one has

$$\begin{aligned} \mathbb{P}\{Y_{j_1} = 1\} &\leq \mathbb{P}\{Y_{j_1} = 1 \mid Y_i = 1\} \\ &\leq \mathbb{P}\{Y_{j_1} = 1 \mid Y_i = 1, Y_{j_2} = 1\} \leq \dots \end{aligned} \quad (85)$$

Proof: Let $Y_{ik} = T\left(\mathbf{M}_{\mathcal{V}_i^{(1)} \cap \mathcal{V}_k^{(1)}}\right)$. Note that if $|\mathcal{V}_i^{(1)} \cap \mathcal{V}_k^{(1)}| \leq 1$, one has $Y_{ik} = 0$.

$$\begin{aligned} \mathbb{P}\{Y_k = 1 \mid Y_i = 1\} &= \sum_{y=0}^1 \mathbb{P}\{Y_k = 1, Y_{ik} = y \mid Y_i = 1\} \\ &\stackrel{(a)}{=} \mathbb{P}\{Y_{ik} = 1 \mid Y_i = 1\} + \mathbb{P}\{Y_k = 1, Y_{ik} = 0 \mid Y_i = 1\} \\ &\stackrel{(b)}{=} \mathbb{P}\{Y_{ik} = 1 \mid Y_i = 1\} \\ &\quad + \mathbb{P}\{Y_{ik} = 0 \mid Y_i = 1\} \cdot \mathbb{P}\{Y_k = 1 \mid Y_{ik} = 0\} \\ &= \mathbb{P}\{Y_{ik} = 1 \mid Y_i = 1\} \cdot (1 - \mathbb{P}\{Y_k = 1 \mid Y_{ik} = 0\}) \\ &\quad + \mathbb{P}\{Y_k = 1 \mid Y_{ik} = 0\}, \end{aligned} \quad (86)$$

where (a) comes from Property 2 which states that if $Y_{ik} = 1$, then $Y_k = 1$. Then (b) is by

$$\begin{aligned} &\mathbb{P}\{Y_k = 1, Y_{ik} = 0 \mid Y_i = 1\} \\ &= \mathbb{P}\{Y_{ik} = 0 \mid Y_i = 1\} \cdot \mathbb{P}\{Y_k = 1 \mid Y_{ik} = 0, Y_i = 1\} \\ &= \mathbb{P}\{Y_{ik} = 0 \mid Y_i = 1\} \cdot \mathbb{P}\{Y_k = 1 \mid Y_{ik} = 0\} \end{aligned}$$

as Y_i and Y_k are independent knowing that $Y_{ik} = 0$. Similarly,

$$\begin{aligned} \mathbb{P}\{Y_k = 1\} &= \mathbb{P}\{Y_{ik} = 1\} \cdot (1 - \mathbb{P}\{Y_k = 1 \mid Y_{ik} = 0\}) \\ &\quad + \mathbb{P}\{Y_k = 1 \mid Y_{ik} = 0\}. \end{aligned} \quad (87)$$

Moreover,

$$\begin{aligned} &\mathbb{P}\{Y_{ik} = 1 \mid Y_i = 1\} \\ &= \frac{\mathbb{P}\{Y_{ik} = 1, Y_i = 1\}}{\mathbb{P}\{Y_i = 1\}} = \frac{\mathbb{P}\{Y_{ik} = 1\}}{\mathbb{P}\{Y_i = 1\}} \geq \mathbb{P}\{Y_{ik} = 1\}. \end{aligned} \quad (88)$$

>From (86), (87), and (88), one obtains that $\mathbb{P}\{Y_k = 1\} \leq \mathbb{P}\{Y_k = 1 \mid Y_i = 1\}$. The other inequalities of Lemma 11 are proved in the same way. ■

H. Proof of Lemma 10

Consider the reference Node i and a node with random index $K \in \mathcal{B}_i \setminus \{i\}$, both known with non-defective sensors. The location of Node K is uniformly distributed on the disk of center i and radius R_0 . One has to evaluate

$$\begin{aligned} &\mathbb{P}\{Y_K = 1 \mid Y_i = 1, \theta_i = \theta_K = 0, K \in \mathcal{B}_i \setminus \{i\}\} \\ &= \mathbb{P}\left\{T\left(\mathbf{M}_{\mathcal{V}_K^{(1)}}\right) = 1 \mid Y_i = 1, \theta_i = \theta_K = 0, K \in \mathcal{B}_i \setminus \{i\}\right\}. \end{aligned}$$

Since $Y_i = 1$, for any realization $\mathbf{m}_{\mathcal{B}_i}$, there exists at least a pair of nodes $(j_1, j_2) \in \mathcal{B}_i^2$ with $j_1 \neq j_2$ such that $[m_{j_1}] \cap [m_{j_2}] = \emptyset$. For the random measurement vector $\mathbf{M}_{\mathcal{B}_i}$, let \mathcal{D}_i the set of such (random) pairs of node indexes. If for some $(J_1, J_2) \in \mathcal{D}_i$, one has also $(J_1, J_2) \in \left(\mathcal{V}_K^{(1)}\right)^2$, then $T\left(\mathbf{M}_{\mathcal{V}_K^{(1)}}\right) = 1$. As a consequence,

$$\begin{aligned} & \mathbb{P}\left\{T\left(\mathbf{M}_{\mathcal{V}_K^{(1)}}\right) = 1 \mid Y_i = 1, \theta_i = \theta_K = 0, K \in \mathcal{B}_i \setminus \{i\}\right\} \\ & \geq \mathbb{P}\left\{J_1 \in \mathcal{V}_K^{(1)}, J_2 \in \mathcal{V}_K^{(1)} \mid (J_1, J_2) \in \mathcal{D}_i, K \in \mathcal{B}_i \setminus \{i\}\right\}. \end{aligned}$$

One has thus to evaluate the probability that $\mathcal{V}_K^{(1)}$ contains the nodes J_1 and J_2 , which themselves belong to \mathcal{B}_i . To evaluate this probability, the fact that (J_1, J_2) belong to \mathcal{D}_i is not important, since the outcome of the LODT does not account for the indexes of the nodes. As a consequence

$$\begin{aligned} & \mathbb{P}\left\{T\left(\mathbf{M}_{\mathcal{V}_K^{(1)}}\right) = 1 \mid Y_i = 1, \theta_i = \theta_K = 0, K \in \mathcal{B}_i \setminus \{i\}\right\} \\ & \geq \mathbb{P}\left\{J_1, J_2 \in \mathcal{V}_K^{(1)} \cap \mathcal{B}_i \mid (J_1, J_2) \in \mathcal{D}_i, K \in \mathcal{B}_i \setminus \{i\}\right\} \\ & = \left(\mathbb{P}\left\{J_1 \in \mathcal{V}_K^{(1)} \cap \mathcal{B}_i \mid J_1 \in \mathcal{B}_i, K \in \mathcal{B}_i \setminus \{i\}\right\}\right)^2 \end{aligned} \quad (89)$$

since the locations of the nodes are independent.

The distance $R_{i,K}$ between Nodes i and K is a random variable with distribution

$$\pi_R(r) = 2r/R_0^2. \quad (90)$$

Now, let S_1 be the intersection of the two disks of radius R_0 and centers i and K respectively, see Figure 15. The probability of a node known to belong to \mathcal{N}_i to be in $\mathcal{U}_i \cap \mathcal{U}_K$ is equal to the area of S_1 divided by πR_0^2 , which is a function of $R_{i,K}$. When $R_{i,K} = r$, the area of S_1 is $S_1(\beta) = R_0^2(2\beta - \sin(2\beta))$, with $\beta = \arccos\left(\frac{r}{2R_0}\right)$. Thus, one has to average (89) over all possible values of $R_{i,K}$ to obtain from (90),

$$\begin{aligned} & \mathbb{P}\{Y_K = 1 \mid Y_i = 1, \theta_i = \theta_K = 0, K \in \mathcal{N}_i\} \\ & \geq \int_0^{R_0} \frac{2r}{R_0^2} \left(\frac{S_1(r)}{\pi R_0^2}\right)^2 dr \\ & = \int_{\pi/3}^{\pi/2} 4 \sin(2\beta) \cdot \left(\frac{2\beta - \sin(2\beta)}{\pi}\right)^2 d\beta \\ & = 1 - \frac{\sqrt{3}}{\pi} - \frac{5}{6\pi^2}. \end{aligned} \quad (91)$$

REFERENCES

- [1] W. Li, F. Bassi, D. Dardari, M. Kieffer, and G. Pasolini, "Low-complexity distributed fault detection for wireless sensor networks," in *Proc. IEEE ICC*, London, United Kingdom, Jun. 2015.
- [2] R. Verdone, D. Dardari, G. Mazzini, and A. Conti, *Wireless sensor and actuator networks: technologies, analysis and design*. Academic Press, 2010.
- [3] V. Barnett and T. Lewis, *Outliers in statistical data*. Wiley New York, 1994, vol. 3.
- [4] Y. Zhang, N. Meratnia, and P. Havinga, "Outlier detection techniques for wireless sensor networks: A survey," *IEEE Communications Surveys & Tutorials*, vol. 12, no. 2, pp. 159–170, 2010.
- [5] A. Mahapatro and P. M. Khilar, "Fault diagnosis in wireless sensor networks: A survey," *Communications Surveys & Tutorials, IEEE*, vol. 15, no. 4, pp. 2000–2026, 2013.

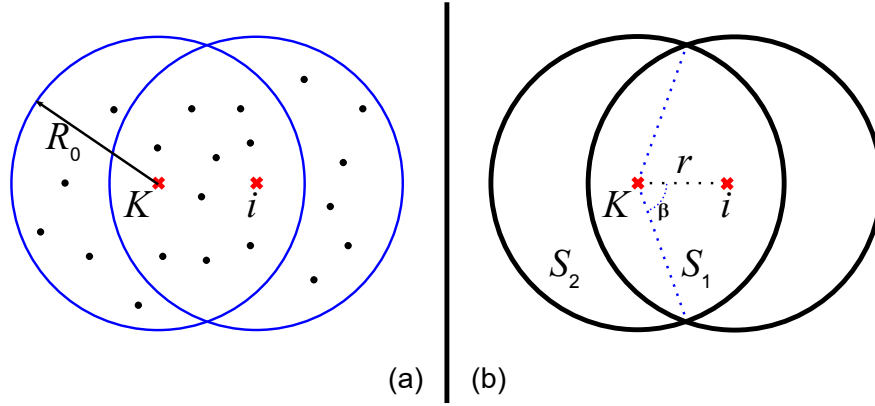


Fig. 15. Sensor configuration.

- [6] H. Dong, Z. Wang, S. X. Ding, and H. Gao, "A survey on distributed filtering and fault detection for sensor networks," *Mathematical Problems in Engineering*, 2014.
- [7] J. Chen, S. Kher, and A. Somani, "Distributed fault detection of wireless sensor networks," in *Proc Workshop DIWANS*, New York, NY, 2006, pp. 65 – 72.
- [8] A. De Paola, S. Gaglio, G. Re, F. Milazzo, and M. Ortolani, "Adaptive distributed outlier detection for wsns," *Cybernetics, IEEE Transactions on*, vol. 45, no. 5, pp. 888–899, May 2015.
- [9] J. W. Branch, C. Giannella, B. Szymanski, R. Wolff, and H. Kargupta, "In-network outlier detection in wireless sensor networks," *Knowledge and information systems*, vol. 34, no. 1, pp. 23–54, 2013.
- [10] R. Isermann, "Model-based fault-detection and diagnosis—status and applications," *Annual Reviews in control*, vol. 29, no. 1, pp. 71–85, 2005.
- [11] C. Lo, M. Liu, J. P. Lynch, and A. C. Gilbert, "Efficient sensor fault detection using combinatorial group testing," in *IEEE Int. Conf. on DCOSS*, 2013, pp. 199–206.
- [12] M. Ding, D. Chen, K. Xing, and X. Cheng, "Localized fault-tolerant event boundary detection in sensor networks," in *Proc IEEE INFOCOM*, Miami, FL, 2005, pp. 902 – 913.
- [13] M. Panda and P. Khilar, "Distributed self fault diagnosis algorithm for large scale wireless sensor networks using modified three sigma edit test," *Ad Hoc Networks*, vol. 25, pp. 170–184, 2015.
- [14] J.-L. Gao, Y.-J. Xu, and X.-W. Li, "Weighted-median based distributed fault detection for wireless sensor networks," *Journal of Software*, vol. 18, no. 5, pp. 1208 – 1217, 2007.
- [15] S. Ji, S.-F. Yuan, T.-H. Ma, and C. Tan, "Distributed fault detection for wireless sensor based on weighted average," in *Proc NSWCTC*, Wuhan, China, 2010, pp. 57 – 60.
- [16] M.-H. Lee and Y.-H. Choi, "Fault detection of wireless sensor networks," *Computer Communications*, vol. 31, no. 14, pp. 3469–3475, 2008.
- [17] J.-Y. Choi, S.-J. Yim, Y.-J. Huh, and Y.-H. Choi, "A distributed adaptive scheme for detecting faults in wireless sensor networks," *WSEAS Transactions on Communications*, vol. 8, no. 2, pp. 269–278, 2009.
- [18] D.-Z. Du and F. K. Hwang, *Combinatorial group testing and its applications*. World Scientific, 1993.
- [19] M. Cheraghchi, A. Hormati, A. Karbasi, and M. Vetterli, "Group testing with probabilistic tests: theory, design and application," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 7057 – 7067, 2011.
- [20] T. Tomic, N. Thomos, and P. Frossard, "Distributed sensor failure detection in sensor networks," *Signal Processing*, vol. 93, no. 2, pp. 399–410, 2013.
- [21] M. Milanese, J. Norton, H. Piet-Lahanier, and E. Walter, Eds., *Bounding Approaches to System Identification*. New York, NY: Plenum Press, 1996.
- [22] L. Jaulin, M. Kieffer, O. Didrit, and E. Walter, *Applied Interval Analysis*. London: Springer-Verlag, 2001.
- [23] M. Haenggi, *Stochastic geometry for wireless networks*. Cambridge University Press, 2012.

- [24] M. Haenggi, M. G. Andrews, F. Baccelli, O. Dousse, and M. F. Franceschetti, “Stochastic geometry and random graphs for the analysis and design of wireless networks,” *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 7, pp. 1029–1046, 2009.
- [25] W. Rudin, *Principles of mathematical analysis*. McGraw-Hill New York, 1964, vol. 3.